

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:22:32 ; Search time 31.5 Seconds  
(without alignments)  
3375.248 Million cell updates/sec

Title: US-09-509-994-1

Perfect score: 2916

Sequence: 1 MGVLVGLALAGLGFAP.....PSPTPGSTLTPPAVLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriapi.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2654	91.0	468	Q9UC32	Q9uc32 homo sapien
2	1840	63.1	577	O35370	O35370 rattus norv
3	1396.5	47.9	461	P97883	P97883 rattus norv
4	580.5	19.9	757	Q9HCU0	Q9hcu0 homo sapien
5	575	19.7	765	O912V1	O912v1 mus musculus
6	575	19.7	765	O91V98	O91v98 mus musculus
7	433	14.8	1664	O9TVQ2	O9tvq2 caenorhabdi
8	414	14.2	1574	O88281	O88281 rattus norv
9	385	13.2	1394	O9V589	O9v589 drosophila
10	375	12.9	708	P87363	P87363 gallus gall
11	374	12.8	2906	O9WUH9	O9wuh9 rattus norv
12	369	12.7	1246	O75095	O75095 homo sapien
13	366	12.6	3857	O88840	O88840 mus musculus
14	365	12.5	2809	Q96JP8	Q96jp8 homo sapien
15	360	12.3	528	O9CXD8	O9cxd8 mus musculus
16	360	12.3	2872	O9WUH8	O9wuh8 rattus norv

17	359	12.3	937	4	Q96FT5	Q96ft5 homo sapien
18	358.5	12.3	1174	11	Q99K58	Q99k58 mus musculus
19	353.5	12.1	746	4	Q96HB9	Q96hb9 homo sapien
20	353.5	12.1	1256	4	Q9NS15	Q9ns15 homo sapien
21	353.5	12.1	1382	4	Q9H7K2	Q9h7k2 homo sapien
22	352.5	12.1	741	4	Q96K89	Q96k89 homo sapien
23	351.5	12.1	1511	4	O75412	O75412 homo sapien
24	351.5	12.1	1587	4	O00508	O00508 homo sapien
25	343	11.8	5636	4	Q96RW7	Q96rw7 homo sapien
26	341	11.7	576	4	Q9Y3V7	Q9y3v7 homo sapien
27	338	11.6	1062	11	Q60789	Q60789 mus musculus
28	337.5	11.6	1289	5	Q8SSS3	Q8sss3 dictyosteli
29	337	11.6	2673	4	Q96SC3	Q96sc3 homo sapien
30	333.5	11.4	961	11	O9EOC6	O9eqc6 mus musculus
31	333	11.4	999	4	Q9NQ36	Q9nq36 homo sapien
32	331.5	11.4	956	11	Q8R542	Q8r542 mus musculus
33	331	11.4	589	5	Q9TZS1	Q9tzs1 caenorhabdi
34	330.5	11.3	956	11	Q99K64	Q99k64 mus musculus
35	330	11.3	2189	5	Q9BI05	Q9bi05 elmeria ten
36	328.5	11.3	1713	11	O88349	O88349 mus musculus
37	328.5	11.3	1764	11	O35806	O35806 rattus norv
38	327.5	11.2	1095	11	Q60784	Q60784 mus musculus
39	323	11.1	1253	11	Q61810	Q61810 mus musculus
40	320.5	11.0	937	5	Q9BLJ1	Q9blj1 ciona intes
41	320	11.0	1394	4	Q8TD95	Q8td95 homo sapien
42	319.5	11.0	1833	11	O08999	O08999 mus musculus
43	315.5	10.8	443	4	Q96TF5	Q96tf5 homo sapien
44	315	10.8	683	4	Q8TBH8	Q8tbh8 homo sapien
45	315	10.8	997	11	Q9JJ50	Q9jj50 mus musculus

#### ALIGNMENTS

#### RESULT 1

Q9UC32 PRELIMINARY; PRT; 468 AA.  
 ID Q9UC32  
 AC Q9UC32;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE THROMBOMODULIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93293792; PubMed=8390446;  
 RA Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;  
 RT "Urinary thrombomodulin, its isolation and characterization.";  
 RL J. Biochem. 113:433-440(1993).  
 DR HSSP; P07204; 1ZAQ.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001304; Lectin\_C.  
 DR InterPro; IPR001491; Thrombomoduln.  
 DR Pfam; PF00008; EGF; 4.  
 DR PRINTS; PR00059; lectin\_c; 1.  
 DR SMART; SM00034; CLECT; 1.  
 DR SMART; SM00181; EGF; 6.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00041; C\_TYPE\_LECTIN\_2; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 SQ SEQUENCE 468 AA; 49444 MW; 4BFE8E98EFB6A40 CRC64;

Query Match 91.0%; Score 2654; DB 4; Length 468;  
 Best local similarity 99.6%; Pred. No. 7.2e-205;  
 Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 19 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRHMTVRSVAADVLSLLNGD 78
DB 1 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRHMTVRSVAADVLSLLNGD 60
QY 79 GVGRRRLWIGLQPLPGCGDKPKRLGRFQWVTGDNNTYSRWARLDLNGAPLCPLCV 138
DB 61 GVGRRRLWIGLQPLPGCGDKPKRLGRFQWVTGDNNTYSRWARLDLNGAPLCPLCV 120
QY 139 AVSAAATVPSEPIWEEOQCEVKADGFLCEHFPATCRPLAVPGAAAVSITYTTPFA 198
DB 121 AVSAAATVPSEPIWEEOQCEVKADGFLCEHFPATCRPLAVPGAAAVSITYTTPFA 180
QY 199 ARGADQALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 258
DB 181 ARGADQALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 240
QY 259 IPGAPRCQPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQGSYSCMCTGYRLAA 318
DB 241 IPGAPRCQPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQGSYSCMCTGYRLAA 300
QY 319 DQHRCEVDVDCILEPSPCQRCVNTGGFECCHYPNYDLVDGECVDPDPCFRANCEYQC 378
DB 301 DQHRCEVDVDCILEPSPCQRCVNTGGFECCHYPNYDLVDGECVDPDPCFRANCEYQC 360
QY 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQACPADCDPNTQASCEGPEGYILDDGFI 438
DB 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQACPADCDPNTQASCEGPEGYILDDGFI 420
QY 439 CTDIDECENGFGSGVCHNLPGTFECICGPDPSALVRHIGTDCDSKVD 486
DB 421 CTDIDECENGFGSGVCHNLPGTFECICGPDPSALVRHIGTDCDSKVD 468

RESULT 2
O35370 PRELIMINARY; PRT; 577 AA.
AC O35370;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-136 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022743; AAB80760.1; -.
DR EMBL; AF022742; AAB80923.1; -.
DR HSSP; P07204; 1FGD.
DR InterPro; IPR001152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR001491; Thrbomodulin.
DR Pfam; PF00008; EGF; 5.
DR PRINTS; PR00059; lectin.c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00001; EGF-like; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
```

```
DR PROSITE; PS01187; EGF_CA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 577 AA; 61844 MW; 0BE764C8BF18555F CRC64;

Query Match 63.1%; Score 1840; DB 11; Length 577;
Best Local Similarity 63.8%; Pred. No. 1.8e-139;
Matches 332; Conservative 49; Mismatches 133; Indels 6; Gaps 5;

QY 1 MLGVLVIGLALAGLPGPAPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRHLM 60
DB 1 MLGVLVIGLALAGLPGPAPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDKPKRLGRFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDKPKRLGRFQWVTGDNNTSYS 118
QY 121 RWARLDLNGAPLCPLCVAVSAAATVPSEPIWEEOQCEVKADGFLCEHFPATCRPLAV 180
DB 119 RWARLDLNGAPLCPLCVAVSAAATVPSEPIWEEOQCEVKADGFLCEHFPATCRPLAV 178
QY 181 EP-GAAAAAVSITYTTPFAARGADQALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREA 239
DB 179 NTRDPEGAHSSITYNTPLGVSGADFTLPIGSSATVAPFGLVLCRALPGTSEGHWTREV 238
QY 240 PGAWDCSVENGCEHACNAIPGAPRCQPAGAAQADGRSCTASATQSCNDLCEHFCVNP 299
DB 239 TGAWDCSVENGCEHACNAIPGAPRCQPAGAAQADGRSCTASATQSCNDLCEHFCVNP 298
QY 300 PDQGSYSCMCTGYRLAAQDHRCEVDVDCILEPSPCQRCVNTGGFECCHYPNYDLVD 359
DB 299 SDVPGSYSCMCTGYRLAAQDHRCEVDVDCILEPSPCQRCVNTGGFECCHYPNYDLVD 358
QY 360 GECVDPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQACPADCDP 419
DB 359 GECVDPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQACPADCDP 418
QY 420 NTOASCEPGEYILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDPSALVRHIGT 479
DB 419 NSPFCQCPGEGFILDGSICTDIDECENGFGSGVCHNLPGTFECICGPDPSALVRHIGT 478
QY 480 CDSGKV-DGGDSGSGEPPPS-PTPGSTLTPPAGVLVHSG 516
DB 479 CDPVLEDESDGSGEHPSSNPTVSSVTPPSARPMHSG 518

RESULT 3
P97883 PRELIMINARY; PRT; 461 AA.
AC P97883;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Thrombomodulin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN CAPILLARY;
RA Wang L., Tran N.D., Schreiber S.S., Zlokovic B.V.;
RL Nucleotide sequence of rat thrombomodulin.;
DR EMBL; U90121; AAB49723.1; -.
DR HSSP; P07204; 1FGD.
DR InterPro; IPR001152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001491; Thrbomodulin.
DR Pfam; PF00008; EGF; 5.
DR PRINTS; PR00097; THRBOMODULN.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00001; EGF-like; 3.
```

```
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1
SQ SEQUENCE 461 AA; 49499 MW; 45971BCE84688E67 CRC64;

Query Match 47.9%; Score 1396.5; DB 11; Length 461;
Best Local Similarity 62.4%; Pred. No. 5.2e-104;
Matches 244; Conservative 37; Mismatches 105; Indels 5; Gaps 4;

Qy 131 PLCGPLCVAVSAATVPSEIWEQCEVKADGFLCEFPHPATC-RPLAVER-GAAAAA 188
Db 12 PLCGPLCVTVSTATAPGEPFAWEKFCENETKGLCEFYFAAFCSRLRYNTRDPEGAH 71
Qy 189 VSIYTGTPFAARGADFQALPGVSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVE 248
Db 72 ISSYNTPLGVSGADFQTLPGSSATVAPFGLVLCRALPDTSGHWTREVTGAWNCSVE 131
Qy 249 NGGCEHACNAPIGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVPPNPDPGYSYC 308
Db 132 NGGCEYMCNRSANGPRCVPSCGGDLQADGRSCAKPVGQLCNELCQHEFCVNNSDVPGYSYC 191
Qy 309 MCEYGYRLAADQRCEDVDDCILPSPQRCVNTQGGFCHCYPNTDLDVDCGEVPEVDP 368
Db 192 MCEYGYRLAADQRCEDVDDCKQGNPCPOLCSNTEGFCRCYDGYELVDGECVEQLDP 251
Qy 369 CFRANCEYQOCPLQNTSYLVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECP 428
Db 252 CFRSKECYQOCVNSTHYNCICAGFAPKLDPPDRCEMFCNETSCPADCDPNSFSQCP 311
Qy 429 EGYILDGFICTDIDECENGFCGVCNHLNPTGFCICGPDPSALVRHIGTDCDSGKY--D 486
Db 312 EGYILDGFICTDIDECQSGECLNCRNLPSGYECIGPDPDTALAGQISKDCDPIPVLED 371
Qy 487 GDSGSGGPPPS-PTPGTGLTPPAVLGVHSG 516
Db 372 SEDGSGGHPSSNPTVSTVPPSARPMHSG 402

RESULT 4
Q9HCU0 PRELIMINARY; PRT; 757 AA.
AC Q9HCU0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Tumor endothelial marker 1 precursor (Endostialin protein).
GN TEM1 OR ENDOSIALIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20407466; PubMed=10947989;
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,
RA Montgomery E., Lal A., Riggins G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.;
RT "Genes expressed in human tumor endothelium.";
RL Science 289:1197-1202(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21269274; PubMed=11084048;
RA Christian S., Ahorn H., Koehler A., Eisenhaber F., Rodi H.P.,
RA Garin-Chesa P., Park J.E., Rettig W.J., Lenter M.C.;
RT "Molecular cloning and characterization of Endostialin, a C-type
lectin-like cell surface receptor of tumor Endothelium.";
RL J. Biol. Chem. 276:7408-7414(2001).
DR EMBL; AF279142; AAC00867.1;
DR EMBL; AJ295846; CAC34381.1;
DR HSSP; P07204; IZQA.
DR InterPro; IPR000152; Asx_hydroxyl.
```

```
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF000008; EGF_2.
DR Pfam; PF000059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_CA; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 757 TUMOR ENDOTHELIAL MARKER 1.
SQ SEQUENCE 757 AA; 80858 MW; C96363EA1FD8FFA0 CRC64;

Query Match 19.9%; Score 580.5; DB 4; Length 757;
Best Local Similarity 33.4%; Pred. No. 2.5e-38;
Matches 176; Conservative 44; Mismatches 192; Indels 115; Gaps 24;

Qy 1 MLGVLVLGALALAGLG-FPAPAEPPQGGSCVHEHDCFALYPGPATFLNASQICDGLRGHL 59
Db 2 LLRLLLAWAAGPTLQGDPAAEPR---AACGPSSCTALFPRRRTFLEAWRACRELGGDL 58
Qy 60 MTRVSSVAADVISILLNGDGGVGRRLWIGLQLPQCGDPRKLRGLRGLRGFWVTGDNNTSY 119
Db 59 ATPRTPEAQRVDSLVG--AGPASRLWIGLQRAQRCQLOR--PLRGFTTTGTDQDTAF 114
Qy 120 SRNARLDNGAPLGGPLCVAVSAATVPSEIWEQCEVKADGFLCEFPHPATCRPLA 179
Db 115 TNWQAQ-PASGGPCPAQRCAVALEAS-----GEHRWLEGSLTAVDGYLCQFEGACPALQ 168
Qy 180 VERGAAAAVSIYTGTPFAARGADFQALPGVSSAAV--APLGLQLMCTAPPAGVQGHWA 236
Db 169 DEAGQAGPAV---YTTFFHLVSTEFELWLPFGSVAAVOCQAGRASLVCVQPEGGV-WS 224
Qy 237 REAP---GAWDCSVENGCEHAC-NAIPGAPRCOCAGALQADGRSCTASATQS-CNDL 291
Db 225 RAGPLCLGT-GCSPDNGGCEHCEVEVDGHVSCRCTEGFRLAADGRSCDPCQAQACEQQ 283
Qy 292 CEHFCVNPDPGYSYSCMCTGYRLAADO-HRCEVDVDCILEPSPQRCVNTQGGFCH 350
Db 284 CE-----PGSPQGYSCNRLGFRPAEDDPRHCYVTDDECCI-AGVCQCMCVNYVGGFECY 336
Qy 351 CYPNYDLVDGCEVPDPCFRANCEYQOCPLNQTSYLVCAEGFAPIPHEPHRCOMFCNQ 410
Db 337 CSEGELE-----ADGIS----- 349
Qy 411 TACPADCDP---NTQASCECEGYILDDGFICTDIDE---CENGFCGSGYCHNLPGT-- 461
Db 350 -----CSPAGAMGAQASQDLGD-ELLDDGDEDEDEAKWAFNGWT-----EMPGLW 397
Qy 462 FECICGPDPSALVRHIGTDCDSKVDGSGSGSEPPSPPTPGSTLTPP 508
Db 398 MEPTQPPDFALAYRPSFPED-----REPQIPYEPPTWPPP 432

RESULT 5
Q912V1 PRELIMINARY; PRT; 765 AA.
ID Q912V1
AC Q912V1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Endostialin.
GN TEM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```





[illegible]

QY	446	ENGFGSCVCHNLPGTFCICGPDPSALVRHIGT-----DCDS-----GKVDGDSG	491
DB	393	ENNGDCSQICVNLAGSVECCQKPGFRLMKORKTCEDISECSNNGCEQICSNQEGGYMC	444
QY	492	SCEP 495	
DB	443	SCEP 446	
RESULT 8			
ID	088281	PRELIMINARY; PRT; 1574 AA.	
AC	088281;		
DT	01-NOV-1998	(TrEMBLrel. 08, Created)	
DT	01-NOV-1998	(TrEMBLrel. 08, Last sequence update)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)	
DE	MEGF6.		
GN	MEGF6.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;		
ON	NCBI_TaxID=10116;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;		
RX	MEDLINE=98360089; PubMed=9693030;		
RA	Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;		
RT	"Identification of high-molecular-weight proteins with multiple EGF-		
RT	like motifs by motif-trap screening."		
RL	Genomics 51:27-34(1998).		
DR	EMBL; AB011532; BAA32462.1; -.		
DR	HSSP; P00736; IAPQ.		
DR	InterPro; IPR000152; Asx_hydroxyl.		
DR	InterPro; IPR000561; EGF-like.		
DR	InterPro; IPR001881; EGF_Ca.		
DR	Pfam; PF00008; EGF; 24.		
DR	SMART; SM00179; EGF_Ca; 4.		
DR	SMART; SM00001; EGF-like; 19.		
DR	PROSITE; PS00010; ASX_HYDROXYL; 5.		
DR	PROSITE; PS00022; EGF_1; UNKNOWN_23.		
DR	PROSITE; PS01186; EGF_2; 23.		
DR	PROSITE; PS01187; EGF_Ca; 5.		
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat		
SQ	SEQUENCE 1574 AA; 165445 MW; 2848533DBF77F6E7 CRC64;		
Query Match 14.2%; Score 414; DB 11; Length 1574;			
Best Local Similarity 34.1%; Pred. No. 1.3e-24;			
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps			
QY	224	CTAPPG-AVQGHWAREAPGAWDCSVENGCGEACNAIPGAPRCQCQFAGALQADGSCSTA	282
DB	147	CRCPGQYQLQD-GKTCQDVEDCEAHNGCGCHRCVTPGSLCECPGFRFLHTDGTCL-	204
QY	283	SATQSC---NDLCEHFVCP-----NPDQFSGVSCM-----309	
DB	205	-AISCTLNGGGOHQCVLTQTHRCQCRPQYQLQDGEDRRRCVRRSPCAEGNGGCMHICQ	263
QY	310	-----CETGYRLAADQHRCEDVDCLILEPSPQRQCVNTQGFECHECVNYDL-VD	359
DB	264	ELRGLAHGCHPGYQLAADRKTCEVDCEALGIAQAHCGLNTQGSFKVCVCHAGVELGAD	323
QY	360	G-EG-----VEPVDPCFRAN--CEVQCPNLQTSYLCVCAEGFAPIPHEPHRCOMFC	408
DB	324	GRQCYRTEMLVNSCEAGNGCGSGHGSHTSTGPLCTCPRGY-----ELDQKTCIDID	377
QY	409	---NOTACPADCPNTQA---SCPCPEGYILD-DGFICTDDECENG-GFCSGVCHNLPGCT	461
DB	378	DCANSPCCQAC-ANTPGGYECSCFAGYRLNTDCCGCDVDECAHGCHGCEHHCSNLAS	436
QY	462	FECIC 466	
DB	437	FOFCF 441	

RESULT 9  
Q9VS89 PRELIMINARY; PRT; 1394 AA.  
AC Q9VS89;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE CG7526 protein (Fragment).  
GN CG7526.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foutsier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003558; AAF50538.1; -.  
DR HSSP; P00736; 1APO.  
DR FlyBase; FBgn0035798; CG7526.  
DR InterPro; IPR000152; ASX\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_CA.  
DR InterPro; IPR003410; Hyalin.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR InterPro; IPR001491; Thrbomodulin.  
DR Pfam; PF00008; EGF; 11.  
DR Pfam; PF00084; sushi; 2.  
DR PRINTS; PR00907; THRBOMODULN.  
DR SMART; SM00032; CCP; 2.  
DR SMART; SM00179; EGF\_CA; 9.  
DR SMART; SM00001; EGF\_like; 5.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 8.

DR PROSITE; PS01186; EGF\_2; 9.  
DR PROSITE; PS01187; EGF\_CA; 10.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
FT NON\_TER 1394 1394  
SQ SEQUENCE 1394 AA; 152269 MW; CD29380E3162F68A CRC64;  
  
Query Match 13.2%; Score 385; DB 5; Length 1394;  
Best Local Similarity 32.1%; Pred. No. 2.4e-22;  
Matches 99; Conservative 33; Mismatches 110; Indels 66; Gaps 18;  
  
QY 224 CTRAPGAVOG-----HWAREAPGAW-----DCSVENGGEHACNAIPG 261  
DB 563 CLCPGAYGLDNIHVITLSNSSFTDSTSETSAHTCIDIDECSLANGNSHFCQNEPG 622  
QY 262 APRCQCPAGAAQAQGRSGTASATQSC---NDLCEHFCVNPDPQGSYSQCMCTGYRLAA 318  
DB 623 GFQCACPLGYALSEDNRTC---QDIDELDSNGQCSQLCL---NQPGGFACACETGELTP 677  
QY 319 DQHRCEVDVDDCILEPSPQRCVNTQGGFCHCYPNYDLVDGE--CVEPVDPC---FRAN 373  
DB 678 DGFGCADIDECSDYGNCSIDICINLLGTHACACERGYELAKDKLSCLD-VDECAGLLSGG 736  
QY 374 CEYQCOPLNOT-SYLCVCAEGFAPIPHEPHRCOMFNCQTACPA--DCDPTQASC----- 425  
DB 737 CSHEC--INKAGTFEGCGPLGY--ILNDDGR-----SCSPALVGGPPTORSADGCA 785  
QY 426 -EGPEGYIIDDGFICTDIDEC--ENGFCSGVCHNLPGTFECICGP-----DSALVRHI 476  
DB 786 IECNPGYITGSDDKVDIDECQKONG--CSHRCSNTEGSPKSCPPGYELDSQKTCQDI 844  
QY 477 GTDCDSGK 484  
DB 845 -DECDQDK 851  
  
RESULT 10  
P87363 PRELIMINARY; PRT; 708 AA.  
AC P87363;  
DT 01-MAY-1997 (TRENBLrel. 03, Created)  
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)  
DT 01-MAY-2002 (TRENBLrel. 20, Last annotation update)  
DE Fibrillin-1 (Fragment).  
GN FBN1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20152896; PubMed=10691037;  
RA Zhou G., Price C.E., Rosenquist T.H., Gadson P.F., Godfrey M.;  
RT "Partial cloning and sequencing of chick fibrillin-1 cDNA."  
RL In Vitro Cell. Dev. Biol. Anim. 36:19-25(2000).  
DR EMBL; U88872; AAB48531.1; -.  
DR HSSP; P07204; 2ADX.  
DR InterPro; IPR000152; ASX\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_CA.  
DR InterPro; IPR002212; Fibril-assoc.  
DR Pfam; PF00008; EGF; 13.  
DR Pfam; PF00683; TB; 2.  
DR PRINTS; PR00907; THRBOMODULN.  
DR SMART; SM00179; EGF\_CA; 14.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 13.  
DR PROSITE; PS01186; EGF\_2; 10.  
DR PROSITE; PS01187; EGF\_CA; 13.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
FT NON\_TER 708 708  
SQ SEQUENCE 708 AA; 76164 MW; C247271C1DF73361 CRC64;

Query Match	12.94;	Score	375;	DB	13;	Length	708;
Best Local Similarity	31.84;	pred.	No. 7.1e-22;				
Matches	112;	Conservative	26;	Mismatches	108;	Indels	106;
Gaps							20;
Qy	244	DCSVENGCEHACNNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNP---	300				
Db	223	ECSIMNGGCNFCVTGSEGSYECSCQGFALPDHRTCT-----DIDE--CEDNPIC	272				
Qy	301	-----DOPGSYSCMETGYRLADQHRCEVDVDCILEPSPC-PQRCVNTGGGFECHCYP	353				
Db	273	DGGQCTNIPGGEYRLCYDGFMASEDMKTCVDVNECDLHPNLCSTGCTNKGSTFICHDM	332				
Qy	354	NYDLVDGE--CVEPVDPC--FRANCEYQCPLN-OTSYLVCVAGFA-----PIP	398				
Db	333	GYSGKGGTGTGD-INECEIGAHNCDRHAVCTNIPGSGFKCSGSGWIGNGIKTDLDBCS	391				
Qy	399	HEPHRCMFCNQACPADCDNTQAS--CECPGEGYLLDDGFCITDIDE-----CENG-	448				
Db	392	NGTHKCSPH-----ADC-KNTMGSYRLCKEY-TGDGFTCTDLDBCSENLNLCENGQ	442				
Qy	449	-----GF-----CS-----GYCHNLPGTFCICGDSALV	473				
Db	443	CLNAPGGYRCEDMGFLPSLDGKACEDIDECSLPNICYGTGCHNLPGLFRCECEGVYELD	502				
Qy	474	RHIG-----TDCDSCK-VDGGDGSGEPPPTPGSTLTPPAVLV	513				
Db	503	RSGNGCTDVNECADPTTICSTCVNTAGSYTCECPD-----FELNPRVGVG	550				

RESULT	11
ID	Q9WUH9
AC	PRELIMINARY; PRT; 2906 AA.
DT	01-NOV-1999 (TREMBlrel. 12, Created)
DT	01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE	Fibrillin-2.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;
ON	NCBI_TaxID=10116;
RX	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99350231; PubMed=10419698;
RA	Yang Q., Ota K., Iian Y., Kumar A., Wada J., Kashiwara N., Wallner E.
RA	Kanwar I.S.;
RT	"Cloning of rat fibrillin-2 cDNA and its role in branching morphogenesis of embryonic lung.";
RL	Dev. Biol. 212:229-242(1999).
DR	EMBL; AF135060; AAD34439.1; -.
DR	HSP; P35555; IEMN.
DR	InterPro: IPRO02086; Aldehyde_dehydr.
DR	InterPro: IPRO00152; Asx_hydroxyl.
DR	InterPro: IPRO00561; EGF-like.
DR	InterPro: IPRO01881; EGF_Ca.
DR	InterPro: IPRO01438; EGF-II.
DR	InterPro: IPRO02212; Fibril-assoc.
DR	Pfam; PF00008; EGF; 46.
DR	Pfam; PF00683; TB; 9.
DR	PRINTS; PR00010; EGFBLOOD.
DR	SMART; SM00179; EGF_CA; 42.
DR	SMART; SM00001; EGF_like; 4.
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR	PROSITE; PS00010; ASX_HYDROXYL; 43.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR	PROSITE; PS01186; EGF_2; 36.
DR	PROSITE; PS01187; EGF_CA; 43.
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat
SQ	SEQUENCE 2906 AA; 313371 MW; 9EE64E727044EF58 CRC64;

Query Match 12.8%; Score 374; DB 11; Length 2906;  
Best Local Similarity 32.9%; Pred. NO. 4e-21;  
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

[illegible]

RESULT 12

```

075095
ID O75095 PRELIMINARY; PRT; 1246 AA.
AC O75095;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MEGF6 (Fragment).
GN MEGF6.
OS Homo Sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
FN SEQUENCE FROM N.A.
RP TISSUE=BRAIN;
RC MEDLINE=98360089; Pubmed=9693030;
RX Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RL like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
DR EMBL; AB011539; BAA32467.2; -.
DE
GN NON_TER 1
FT SEQUENCE 1246 AA; 130304 MW; BD8E70A0F6A2CB29 CRC64;
SQ

```

Query Match 12.7%; Score 369; DB 4; Length 1246;  
Best Local Similarity 32.6%; Pred. No. 4e-21;  
Matches 95; Conservative 27; Mismatches 97; Indels 72; Gaps 15;

Qy	238	EAPGAWDCSVENGCEHACNAIFCAPRCQCYPAGAAALQADGRSCTASATQSC----	NDLCEH	294
Dd	70	QRPDVDECRTHNGCQHRCYNTPGSYLCECKPCFRLHTDSRTCL--AINSCALNGGCQH	127	
Qy	295	FCV-----PNDPOGPS-----YSMCSTGYR	315	
Dd	128	HCVOLTTTRHRCQCPGFQLQEDGRHCVRKSPCANRGSCMHRCQVVYRGLARCECHVGQ	187	
Qy	316	LAADHRCEDVDCCILEPSPCQRCVNTQGFECHCYPNVNL-VDG-EC----	VEPVDFC	369
Dd	188	LAADGKACEDDECAAGLAQAUGCINTQSFCVCVCHAGYELGADGRQCYRTMETVNSC	247	
Qy	370	FRAN--CEYOCPLNQTSY--LCVCAEGFAPIPEHPHQMFQ-----NOTTACPADC	418	
Dd	248	EANNGGCSHG----SHTSAGPLTCPRGY-----ELDTORTCIDVDDCADSECCOQVCT	299	
Qy	419	PNTQA-SCECPGYILD-DGFICTDIDECENG-GFCSGVGHNLPGTFEGIC	466	



ID	Q9CXD8	PRELIMINARY;	PRT;	528 AA.
AC	Q9CXDB;			
DT	01-JUN-2001	(TrEMBLrel. 17, Created)		
DD	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)		
DE	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)		
DR	G130401LZORIK	protein.		
GN	G130401LZORIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
NCBI_TaxID=10090;				
[1]				
RN	SEQUENCE FROM N.A.			
RC	SFRALNE=C57BL/6J; TISSUE=THYMUS;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukuishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,			
RA	Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,			
RA	Hayashizaki Y.,			
RT	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
DR	ENBL; AK018073; BAB31061.1; -.			
DR	HSSP; P00736; IAPQ.			
DR	MGSD; MG1:1922990; G130401LZORIK.			
DR	InterPro; IPR000152; Asx_hydroxyl.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	pfam; PF00008; EGF; 6.			
DR	SMART; SM00181; EGF; 8.			
DR	SMART; SM00179; EGF_CA; 6.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 3.			
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.			
DR	PROSITE; PS01186; EGF_2; 5.			
DR	PROSITE; PS01187; EGF_CA; 4.			
SW	EGF-like domain; Glycoprotein; Hydroxylation.			
KW	SEQUENCE 528 AA; 58217 MW; 36D66698169328E CRC64;			
Query Match	12.3%; Score 360; DB 11; Length 528;			
Best Local Similarity	24.6%; Pred. No. 8.2e-21;			
Matches 123; Conservative	57; Mismatches 151; Indels 170; Gaps			
QY	86 LWIGLQLPGCGDKPLRGFGFWVTGDNNTSYSEWRDLNGAPLGCLPCVAVSAABA 145	:	:	:
Db	: :	:	:	:
Db	12 VWAAANWPP--GSALQLQP--GMPNVCREEQTLVLRLSR-----PCAQA 51	:	:	:
QY	146 TVPSEPPIWEQQC-----EVKADGLCFEFHPAT-----CRPLAV----EPCAA 185	:	:	:
Db	: :	:	:	:
Db	52 FIDIIFQW-KQGSGRPWCVGVERIRYIIIRHYVATEHQVFRCPCGWQMDPEGCF 110	:	:	:
QY	186 AAVSVITYGTPTFAAGDAFPGLPVGSSAVALGLQLMCTAPPAGVQG-HWAREAPGAWD 244	:	:	:
Db	: :	:	:	:
Db	111 SSLSSL-GTFHSGRCSQDYT-----RQCICS-----QGFHGPHCYQDINE 150	:	:	:
QY	245 CSVENGGC-EHACNAITPGAPROCPCAGALQADGRSC-----TASATQSC-NDL-- 291	:	:	:
Db	: :	:	:	:
Db	151 CAVDNGGCRRDCNTI-GSYCRCQAGOKLEEDGRGEDVBECVVNGCCQQRCLTG 209	:	:	:
QY	292 -----CEHFVCPNPDPQGVSYSCMCGETGYRLAADQHR 322	:	:	:



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:23:13 ; Search time 13.5 seconds  
(without alignments)  
1124.609 Million cell updates/sec

Title: US-09-509-994-1

Perfect score: 2916

Sequence: 1 MGVVLGALALAGLPAP.....PSPTGSLTPPAVLVHSG 516

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTBUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/PTBUSfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2916	100.0	575	1	US-08-261-206A-59
2	2912	99.9	575	1	US-08-312-870-1
3	2912	99.9	575	6	5466668-6
4	2904	99.6	575	1	US-08-170-290A-54
5	2842	97.5	572	6	5256770-7
6	2830	97.1	498	2	US-08-733-564-2
7	2820	96.7	497	1	US-08-312-870-3
8	2770	95.0	494	1	US-08-014-723-14
9	2770	95.0	494	1	US-08-110-011A-14
10	2768	94.9	494	1	US-08-014-723-16
11	2768	94.9	494	1	US-08-110-011A-16
12	2690	92.2	475	1	US-08-307-444A-1
13	2690	92.2	475	1	US-08-587-389-1
14	2686	92.1	475	1	US-08-307-444A-2
15	2686	92.1	475	1	US-08-587-389-2
16	2684	92.0	476	1	US-08-014-723-1
17	2684	92.0	476	1	US-08-110-011A-1
18	2682	92.0	476	1	US-08-014-723-2
19	2682	92.0	476	1	US-08-014-723-18
20	2682	92.0	476	1	US-08-110-011A-2
21	2682	92.0	476	1	US-08-110-011A-18
22	2596	89.0	456	1	US-08-307-444A-3
23	2596	89.0	456	1	US-08-587-389-3
24	2592	88.9	456	1	US-08-307-444A-4
25	2592	88.9	456	1	US-08-587-389-4
26	2543	87.2	446	1	US-08-307-444A-5
27	2543	87.2	446	1	US-08-587-389-5

28	1617	55.5	275	1	US-08-312-870-7	Sequence 7, Appl
29	1159	39.7	215	1	US-08-312-870-5	Sequence 5, Appl
30	685	23.5	114	2	US-08-733-564-1	Sequence 1, Appl
31	685	23.5	115	1	US-08-312-870-9	Sequence 9, Appl
32	580	19.9	652	2	US-08-751-305-2	Sequence 2, Appl
33	573	19.7	492	4	US-09-724-864-39	Sequence 39, Appl
34	358	12.3	638	2	US-08-897-443-1	Sequence 1, Appl
35	352	12.1	58	1	US-08-261-206A-3	Sequence 3, Appl
36	331.5	11.4	1964	4	US-09-467-997-1	Sequence 1, Appl
37	328.5	11.3	956	2	US-08-897-443-3	Sequence 3, Appl
38	323	11.1	1253	3	US-08-479-722B-4	Sequence 4, Appl
39	320	11.0	1394	6	5177197-30	Patent No. 5177197
40	319.5	11.0	1833	3	US-08-479-722B-2	Sequence 2, Appl
41	319.5	11.0	1833	5	PCT-US95-02251-18	Sequence 18, Appl
42	315.5	10.8	443	2	US-08-833-963C-2	Sequence 2, Appl
43	315.5	10.8	443	3	US-08-980-514-1	Sequence 1, Appl
44	308	10.6	448	2	US-08-884-072-1	Sequence 1, Appl
45	308	10.6	448	4	US-09-212-168-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-08-261-206A-59  
; Sequence 59, Application US/08261206A  
; Patent No. 5574007  
; GENERAL INFORMATION:  
; APPLICANT: Zushi, Mitichitaka  
; APPLICANT: Gomi, Komakazu  
; APPLICANT: Yamamoto, Shuji  
; APPLICANT: Suzuki, Koji  
; APPLICANT: Matsuda, Akio  
; TITLE OF INVENTION: A Polypeptide Capable of Interacting  
; TITLE OF INVENTION: with Thrombin  
; NUMBER OF SEQUENCES: 80  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: 301 N. Washington St.  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22046-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/261.206A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/740,492  
; FILING DATE: 03-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svensson, Leonard R.  
; REGISTRATION NUMBER: 30330  
; REFERENCE/DOCKET NUMBER: 216-275P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-241-1300  
; TELEFAX: 703-241-2848  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 575 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:

*NO HOMOLOGY*



```
; NAME/KEY: Protein
; LOCATION: 1..575
; OTHER INFORMATION: /label= protein
; OTHER INFORMATION: /note= "human thrombomodulin"
US-08-261-206A-59

Query Match      100.0%; Score 2916; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 1.2e-198;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEPQPGSGQCVHDCFALYPGPATFELNASQICDGLRGHLM 60
Db 1 MLGVLVGLALAGLGFPAEPQPGSGQCVHDCFALYPGPATFELNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCGDKRGLGRLGFWQVTTGDNNTSYS 120
Db 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCGDKRGLGRLGFWQVTTGDNNTSYS 120
QY 121 RWRDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAV 180
Db 121 RWRDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAV 180
QY 181 EPGAAAAYSIITYGTFFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAAYSIITYGTFFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGGECHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGGECHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECHECPNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECHECPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCTACPADCDPN 420
QY 421 TQASCEPEGYILLDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
Db 421 TQASCEPEGYILLDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGSGSGGEPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSGSGGEPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 2
US-08-312-870-1
; Sequence 1, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmen, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:
; CLASSIFICATION: 435
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 19..575
US-08-312-870-1

Query Match      99.9%; Score 2912; DB 1; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.3e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEPQPGSGQCVHDCFALYPGPATFELNASQICDGLRGHLM 60
Db 1 MLGVLVGLALAGLGFPAEPQPGSGQCVHDCFALYPGPATFELNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCGDKRGLGRLGFWQVTTGDNNTSYS 120
Db 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCGDKRGLGRLGFWQVTTGDNNTSYS 120
QY 121 RWRDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAV 180
Db 121 RWRDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAV 180
QY 181 EPGAAAAYSIITYGTFFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAAYSIITYGTFFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGGECHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGGECHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECHECPNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECHECPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCTACPADCDPN 420
QY 421 TQASCEPEGYILLDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
Db 421 TQASCEPEGYILLDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGSGSGGEPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSGSGGEPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 3
5466668-6
; Patent No. 5466668
; APPLICANT: GLASER, CHARLES B.;MORSER, MICHAEL J.;LIGHT,
; DAVID R.
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR
; PHARMACEUTICAL USE
; NUMBER OF SEQUENCES: 57
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,346
; FILING DATE: 22-NOV-1993
```

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 568,456  
; FILING DATE: 15-AUG-1990  
; APPLICATION NUMBER: 506,325  
; FILING DATE: 09-APR-1990  
; APPLICATION NUMBER: 406,941  
; FILING DATE: 13-SEP-1989  
; APPLICATION NUMBER: 345,374  
; FILING DATE: 28-APR-1989  
; SEQ ID NO: 6  
; LENGTH: 575  
5466668-6

Query Match 99.9%; Score 2912; DB 6; Length 575;  
Best Local Similarity 99.8%; Pred. No. 2.3e-198;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVLGALAGLGGFPAPAPQPGSQCVHEHDFALYPGPATFLNASQICDGLRHLM 60  
DB 1 MGVLVLGALAGLGGFPAPAPQPGSQCVHEHDFALYPGPATFLNASQICDGLRHLM 60

QY 61 TVRSSVAADVITSLINGDGGVRRRLWTGLQPPCGGDPKRLGRLGFWTGDNTSYS 120  
DB 61 TVRSSVAADVITSLINGDGGVRRRLWTGLQPPCGGDPKRLGRLGFWTGDNTSYS 120

QY 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEQQCEVKADGFLCFEHPFATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEQQCEVKADGFLCFEHPFATCRPLAV 180

QY 181 EFGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240  
DB 181 EFGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240

QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300

QY 301 DPGSYSCMCTGYRLAADHRCEDVDDCILEPSPCPCORCVNTQGGFCHCYPNYDLVDG 360  
DB 301 DPGSYSCMCTGYRLAADHRCEDVDDCILEPSPCPCORCVNTQGGFCHCYPNYDLVDG 360

QY 361 ECVEPVDFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
DB 361 ECVEPVDFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCCHNLPGTFECICGPDLSALRHIGTDC 480  
DB 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCCHNLPGTFECICGPDLSALRHIGTDC 480

QY 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516  
DB 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516

RESULT 4  
US-08-170-290A-54  
; Sequence 54, Application US/08170290A  
; Patent No. 5702931  
; GENERAL INFORMATION:  
; APPLICANT: Andrews, William H.  
; APPLICANT: Morser, Michael J.  
; APPLICANT: Zieglender, Laura R.  
; TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and  
; TITLE OF INVENTION: Compositions  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James M. Heslin  
; STREET: 379 Lytton Ave.  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/170,290A  
; FILING DATE: 28-DEC-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05573  
; FILING DATE: 01-JUL-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/724,237  
; FILING DATE: 01-JUL-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Heslin, James M.  
; REGISTRATION NUMBER: 29,541  
; REFERENCE/DOCKET NUMBER: 11973-58-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 575 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-170-290A-54

Query Match 99.6%; Score 2904; DB 1; Length 575;  
Best Local Similarity 99.6%; Pred. No. 8.4e-198;  
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGVLVLGALAGLGGFPAPAPQPGSQCVHEHDFALYPGPATFLNASQICDGLRHLM 60  
DB 1 MGVLVLGALAGLGGFPAPAPQPGSQCVHEHDFALYPGPATFLNASQICDGLRHLM 60

QY 61 TVRSSVAADVITSLINGDGGVRRRLWTGLQPPCGGDPKRLGRLGFWTGDNTSYS 120  
DB 61 TVRSSVAADVITSLINGDGGVRRRLWTGLQPPCGGDPKRLGRLGFWTGDNTSYS 120

QY 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEQQCEVKADGFLCFEHPFATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEQQCEVKADGFLCFEHPFATCRPLAV 180

QY 181 EFGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240  
DB 181 EFGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240

QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300

QY 301 DPGSYSCMCTGYRLAADHRCEDVDDCILEPSPCPCORCVNTQGGFCHCYPNYDLVDG 360  
DB 301 DPGSYSCMCTGYRLAADHRCEDVDDCILEPSPCPCORCVNTQGGFCHCYPNYDLVDG 360

QY 361 ECVEPVDFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
DB 361 ECVEPVDFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCCHNLPGTFECICGPDLSALRHIGTDC 480  
DB 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCCHNLPGTFECICGPDLSALRHIGTDC 480

QY 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516  
DB 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516

RESULT 5

```
5256770-7
; Patent No. 5256770
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT, DAVID R.
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/506,325
; FILING DATE: 09-APR-1990
; SEQ ID NO: 7:
; LENGTH: 572
5256770-7

Query Match          97.5%; Score 2842.5; DB 6; Length 572;
Best Local Similarity 98.4%; Pred. No. 1.8e-193;
Matches 508; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 1 MLGVLVGLALAGLGFPAEPAPQPGSQCVHEHDCFALYGPATFLNASQICDGLRGLHLM 60
DB 1 MLGVLVGLALAGLGFPAEPAPQPGSQCVHEHDCFALYGPATFLNASQICDGLRGLHLM 60
QY 61 TVRSSVAADVLSLLNGDGVGRRRLWIGLQLPPGCGDKRRLGRLGFWQVWTDGNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGVG--RRLWIGLQLPPGCGDKRRLGRLGFWQVWTDGNTSYS 119
QY 121 RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFHFPATCRPLAV 179
QY 181 EPGAAAASVITGTPTFAAGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 180 EPGAAAASVITGTPTFAAGADFOALPVGSSAAVAPLGLQMLCTA--GNVQGHWAREAP 237
QY 241 GAWDCSVENGCGCEHACNAIFAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNP 300
DB 238 GAWDCSVENGCGCEHACNAIFAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNP 297
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360
DB 298 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 357
QY 361 ECVEPVPDPCFRANCEYOCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
DB 358 ECVEPVPDPCFRANCEYOCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 417
QY 421 TQASCECEGYILLDGFICTDIDCEENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
DB 418 TQASCECEGYILLDGFICTDIDCEENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 477
QY 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVGLVHSG 516
DB 478 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVGLVHSG 513
```

```
RESULT 6
US-08-733-564-2
; Sequence 2, Application US/08733564
; Patent No. 5916874
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, Kenji
; APPLICANT: MOCHIDA, Satoshi
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/733,564
; FILING DATE: 18 OCTOBER 1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0216-0362P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-733-564-2

Query Match          97.1%; Score 2830; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.2e-192;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAEPQPGSQCVHEHDCFALYGPATFLNASQICDGLRGLHLMTVRSSVAADVLSLLNGD 78
DB 1 APAEPQPGSQCVHEHDCFALYGPATFLNASQICDGLRGLHLMTVRSSVAADVLSLLNGD 60
QY 79 GVGRRRLWIGLQLPPGCGDKRRLGRLGFWQVWTDGNTSYSRWARLDLNGAPLCPLCV 138
DB 61 GVGRRRLWIGLQLPPGCGDKRRLGRLGFWQVWTDGNTSYSRWARLDLNGAPLCPLCV 120
QY 139 AVSAAEATVPSEPIWEEQOCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIYGTPTFA 198
DB 121 AVSAAEATVPSEPIWEEQOCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIYGTPTFA 180
QY 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGCEHACNA 258
DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGCEHACNA 240
QY 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNPDPGSGYSCMCTGYRLAA 318
DB 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNPDPGSGYSCMCTGYRLAA 300
QY 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDGECVEPVPDPCFRANCEYOC 378
DB 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDGECVEPVPDPCFRANCEYOC 360
QY 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDNTQASCECPGEGYILLDGF 438
DB 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDNTQASCECPGEGYILLDGF 420
QY 439 CTDIDCEENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDQSGKVDGSDSGSGPPSP 498
DB 421 CTDIDCEENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDQSGKVDGSDSGSGPPSP 480
QY 499 PTPGSTLTTPPAVGLVHSG 516
DB 481 PTPGSTLTTPPAVGLVHSG 498

RESULT 7
US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
```

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Richards, Medlock & Andrews  
STREET: 1201 Elm Street, Suite 4500  
CITY: Dallas  
STATE: Texas  
COUNTRY: US  
ZIP: 75270-2197  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/312,870  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hansen, Eugenia S.  
REGISTRATION NUMBER: 31,966  
REFERENCE/DOCKET NUMBER: OMR# B35150  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 214-939-4500  
TELEFAX: 214-939-4600  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-312-870-3

Query Match	96.7%	Score 2820;	DB 1;	Length 497;
Best Local Similarity	99.8%	Pred. No. 6.2e-192;		
Matches 496;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	19	APAEPOGGSOCVHEHDCFALYPGPATFLNASQICDGLRGHLMTVRSVVAADVISILLNGD	78	
DB	1	APAEPOGGSOCVHEHDCFALYPGPATFLNASQICDGLRGHLMTVRSVVAADVISILLNGD	60	
QY	79	GGVRRRLWTGLQLPPCGGDPKRLGPLRGFWVTGDNNNTSYSRWARDLNGAPLCGPCLCV	138	
DB	61	GGVRRRLWTGLQLPPCGGDPKRLGPLRGFWVTGDNNNTSYSRWARDLNGAPLCGPCLCV	120	
QY	139	AVSAEAETVSEPITWESQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYTGTPFA	198	
DB	121	AVSAEAETVSEPITWESQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYTGTPFA	180	
QY	199	ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAPGAWDCSVENGCGEHACNA	258	
DB	181	ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAPGAWDCSVENGCGEHACNA	240	
QY	259	IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA	318	
DB	241	IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA	300	
QY	319	DQHRCEVDVDCILPSPSCPQRCVNTQGGFCHCYPNTDLVDGCEVPEVPDPCFRANCEYQC	378	
DB	301	DQHRCEVDVDCILPSPSCPQRCVNTQGGFCHCYPNTDLVDGCEVPEVPDPCFRANCEYQC	360	
QY	379	QPLNQTSTYLCVACGEAFPIPEPHRCMFCNQTCAPADCDPNTQASCEPCPGYIILDDGFI	438	
DB	361	QPLNQTSTYLCVACGEAFPIPEPHRCMFCNQTCAPADCDPNTQASCEPCPGYIILDDGFI	420	
QY	439	CTDIDECENGFGCSGVCHNLPGTEFCICGPDPSALVRHIGTDCDQSKVDGSGSGSEPPPS	498	
DB	421	CTDIDECENGFGCSGVCHNLPGTEFCICGPDPSALRHIGTDCDQSKVDGSGSGSEPPPS	480	
QY	499	PTPGSTLTPPAVGLVHS	515	
DB	481	PTPGSTLTPPAVGLVHS	497	

## RESULT 8

US-08-014-723-14  
; Sequence 14, Application US/08014723  
; Patent No. 5273962  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBION, P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,723  
; FILING DATE: 19930208  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Obion, No. 5273962man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-014-723-14

Query Match	95.0%;	Score 2770;	DB 1;	Length 494;
Best Local Similarity	99.4%;	Pred. No. 2.1e-188;		
Matches 491;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY 1	MGVVLVLGALAGLGGFPAPAPQPGGSGQCV	EHDFCALYPGPATFLNASQICDGLRGHLM	60	
DB 1	MGVVLVLGALAGLGGFPAPAPQPGGSGQCV	EHDFCALYPGPATFLNASQICDGLRGHLM	60	
QY 61	TYRSSVAADVISLLINGDGGVRRRLWTGLQL	PPCGDGPKKLGLRGQFWTGDNTNTSYS	120	
DB 61	TYRSSVAADVISLLINGDGGVRRRLWTGLQL	PPCGDGPKKLGLRGQFWTGDNTNTSYS	120	
QY 121	RWARDLNGAPLCGPLCAVSAAEATVSEPTWE	BQQCEVKADGFLCFEFHPATCRPLAV	180	
DB 121	RWARDLNGAPLCGPLCAVSAAEATVSEPTWE	BQQCEVKADGFLCFEFHPATCRPLAV	180	
QY 181	EPGAAAAVSITYGTPFAARGADFOALPVGSS	AAVAPILGLQLMCTAPPGAVOGHWAREAP	240	
DB 181	EPGAAAAVSITYGTPFAARGADFOALPVGSS	AAVAPILGLQLMCTAPPGAVOGHWAREAP	240	
QY 241	GAWDCSVENGCGEHACNATPGAPRCQCPAGA	LQADGRCSTASATQSCNDLCEHFCVNP	300	
DB 241	GAWDCSVENGCGEHACNATPGAPRCQCPAGA	LQADGRCSTASATQSCNDLCEHFCVNP	300	
QY 301	DOPGYSMCMETGYRLAADQHRCEDDVDCIL	EPSPCQRCVNTQGGFECHCYPNFDLVDG	360	
DB 301	DOPGYSMCMETGYRLAADQHRCEDDVDCIL	EPSPCQRCVNTQGGFECHCYPNFDLVDG	360	
QY 361	ECVEPVDPCFRANCEYOCOPLOTNSYLVCY	CAEGFAPIPHEPHRCOMFCNOTACPADCDPN	420	

Db 361 ECVPEVDPFRANCEYQCQLNQTSLVCVCAEGFAPIPHEPHRCQMFNQTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDDIDECENGFCGSGVCHNLPCTFECICGPDSSALVRRHIGTDC 480  
Db 421 TQASCEPGEYILDDGFICTDDIDECENGFCGSGVCHNLPCTFECICGPDSSALVRRHIGTDC 480  
QY 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDGDSGSGE 494  
RESULT 9  
US-08-110-011A-14  
; Sequence 14, Application US/08110011A  
; Patent No. 5354664  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,011A  
; FILING DATE: 23-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5354664man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-110-011A-14  
Query Match 95.0%; Score 2770; DB 1; Length 494;  
Best Local Similarity 99.4%; Pred. No. 2.1e-188;  
Matches 491; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MLGVLVLGALAGLGPAPAEPPQGSQCQVHDCFAFYFGPATFLNASQICDGLRHL 60  
Db 1 MLGVLVLGALAGLGPAPAEPPQGSQCQVHDCFAFYFGPATFLNASQICDGLRHL 60  
QY 61 TVRSSVAADVLSILLNGDGVGRRLRWIGLQLPFGCGDPRRLGRLGFWQVWTDGNTSYS 120  
Db 61 TVRSSVAADVLSILLNGDGVGRRLRWIGLQLPFGCGDPRRLGRLGFWQVWTDGNTSYS 120  
QY 121 RWARLDLNGAPLCGLPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCGLPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180

QY 181 EPGAAAAVSIYGTTPFAARGADFOALPYGVSSAAVAPLGLQLMCTAPPNAVQGHWAREAP 240  
Db 181 EPGAAAAVSIYGTTPFAARGADFOALPYGVSSAAVAPLGLQLMCTAPPNAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGGHACNAIPGAPRCOCPCAGAAQAQAGRSCTASATOSCNLDCHEFCVNP 300  
Db 241 GAWDCSVENGCGGHACNAIPGAPRCOCPCAGAAQAQAGRSCTASATOSCNLDCHEFCVNP 300  
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECHECYNYDLVDG 360  
Db 301 DQPGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECHECYNYDLVDG 360  
QY 361 ECVPEVDPFRANCEYQCQLNQTSLVCVCAEGFAPIPHEPHRCQMFNQTACPADCDPN 420  
Db 361 ECVPEVDPFRANCEYQCQLNQTSLVCVCAEGFAPIPHEPHRCQMFNQTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDDIDECENGFCGSGVCHNLPCTFECICGPDSSALVRRHIGTDC 480  
Db 421 TQASCEPGEYILDDGFICTDDIDECENGFCGSGVCHNLPCTFECICGPDSSALVRRHIGTDC 480  
QY 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDGDSGSGE 494  
RESULT 10  
US-08-014-723-16  
; Sequence 16, Application US/08014723  
; Patent No. 5273962  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,723  
; FILING DATE: 19930208  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5273962man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-014-723-16  
Query Match 94.9%; Score 2768; DB 1; Length 494;  
Best Local Similarity 99.2%; Pred. No. 2.9e-188;  
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGPAPAEPOPGSGQVEHDCFALYPGPATFNLASOICDGLRGHLM 60  
Db 1 MLGVLVGLALAGLGPAPAEPOPGSGQVEHDCFALYPGPATFNLASOICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFWQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFWQWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
Db 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCSEHFCVPPNP 300  
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCSEHFCVPPNP 300  
QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360  
Db 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360  
QY 361 ECVEPVDPFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCDPN 420  
Db 361 ECVEPVDPFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Db 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
QY 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDDEASGSGD 494

RESULT 11  
US-08-110-011A-16  
; Sequence 16, Application US/08110011A  
; Patent No. 5354664  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, McLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,011A  
; FILING DATE: 23-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5354664man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000

; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-110-011A-16  
  
Query Match 94.9%; Score 2768; DB 1; Length 494;  
Best Local Similarity 99.2%; Pred. No. 2.9e-188;  
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0:  
  
QY 1 MLGVLVGLALAGLGPAPAEPOPGSGQVEHDCFALYPGPATFNLASOICDGLRGHLM 60  
Db 1 MLGVLVGLALAGLGPAPAEPOPGSGQVEHDCFALYPGPATFNLASOICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFWQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFWQWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
Db 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCSEHFCVPPNP 300  
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCSEHFCVPPNP 300  
QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360  
Db 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360  
QY 361 ECVEPVDPFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCDPN 420  
Db 361 ECVEPVDPFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Db 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
QY 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDDEASGSGD 494

RESULT 12  
US-08-307-444A-1  
; Sequence 1, Application US/08307444A  
; Patent No. 5516559  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI  
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIFF & BERRIDGE  
; STREET: P.O. BOX 19928  
; CITY: ALEXANDRIA  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/307,444A  
FILING DATE: 19-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/835,436  
FILING DATE: 26-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: OLIFF, JAMES A.  
REGISTRATION NUMBER: 27,075  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6400  
TELEFAX: (703) 836-2787  
TELEX: 90-1799 PTO ALEX  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-307-444A-1

Query Match 92.2%; Score 2690; DB 1; Length 475;  
Best Local Similarity 100.0%; Pred. No. 9e-183;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLGVLVGLALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVGLALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPGFOWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPGFOWVTGDNNTSYS 120  
QY 121 RWARDLNGAPLCGPLCVAVSAEAATVPSEPIWEEOQCEVKADGFLCFEHPFATCRPLAY 180  
Db 121 RWARDLNGAPLCGPLCVAVSAEAATVPSEPIWEEOQCEVKADGFLCFEHPFATCRPLAY 180  
QY 181 EPGAAAAVSIYGTFFAARGADFOALPVGSSAAVAPLGQLMCTAPPGAVOGHWAREAP 240  
Db 181 EPGAAAAVSIYGTFFAARGADFOALPVGSSAAVAPLGQLMCTAPPGAVOGHWAREAP 240  
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGYSVCMCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
Db 301 DQPGYSVCMCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
QY 361 ECVEPVPDCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420  
Db 361 ECVEPVPDCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420  
QY 421 TQASCECPGYLLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALVRH 475  
Db 421 TQASCECPGYLLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALVRH 475

## RESULT 13

US-08-587-389-1  
Sequence 1, Application US/08587389  
Patent No. 5695964  
GENERAL INFORMATION:  
APPLICANT: NII, ATSUSHI  
APPLICANT: MORISHITA, HIDEAKI  
APPLICANT: UEMURA, AKIO  
APPLICANT: MOCHIDA, EI

TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT  
TITLE OF INVENTION: PRODUCTION THEREOF, AND THERAPEUTIC AGENT (AS AMENDED)  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE  
STREET: P.O. BOX 19928  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/587,389  
FILING DATE: 17-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/307,444  
FILING DATE: 19-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: OLIFF, JAMES A.  
REGISTRATION NUMBER: 27,075  
REFERENCE/DOCKET NUMBER: JAO 27706  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6400  
TELEFAX: (703) 836-2787  
TELEX: 90-1799 PTO ALEX  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-587-389-1

Query Match 92.2%; Score 2690; DB 1; Length 475;  
Best Local Similarity 100.0%; Pred. No. 9e-183;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLGVLVGLALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVGLALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPGFOWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPGFOWVTGDNNTSYS 120  
QY 121 RWARDLNGAPLCGPLCVAVSAEAATVPSEPIWEEOQCEVKADGFLCFEHPFATCRPLAY 180  
Db 121 RWARDLNGAPLCGPLCVAVSAEAATVPSEPIWEEOQCEVKADGFLCFEHPFATCRPLAY 180  
QY 181 EPGAAAAVSIYGTFFAARGADFOALPVGSSAAVAPLGQLMCTAPPGAVOGHWAREAP 240  
Db 181 EPGAAAAVSIYGTFFAARGADFOALPVGSSAAVAPLGQLMCTAPPGAVOGHWAREAP 240  
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGYSVCMCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
Db 301 DQPGYSVCMCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
QY 361 ECVEPVPDCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420  
Db 361 ECVEPVPDCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420



QY 421 TQASCEPEGYILDDGFICTDDECENGFCGSGVCHNLPGTTECICGPDPSALVRH 475  
|||||  
Db 421 TQASCEPEGYILDDGFICTDDECENGFCGSGVCHNLPGTTECICGPDPSALVRH 475

RESULT 14  
US-08-307-444A-2  
; Sequence 2, Application US/08307444A  
; Patent No. 551659  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI  
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIFF & BERRIDGE  
; STREET: P.O. BOX 19928  
; CITY: ALEXANDRIA  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/307,444A  
; FILING DATE: 19-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/835,436  
; FILING DATE: 26-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OLIFF, JAMES A.  
; REGISTRATION NUMBER: 27,075  
; REFERENCE/DOCKET NUMBER: JAO 27706  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6400  
; TELEFAX: (703) 836-2787  
; TELEX: 90-1799 PTO ALEX  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-307-444A-2

Query Match 92.1%; Score 2686; DB 1; Length 475;  
Best Local Similarity 99.8%; Pred. No. 1.7e-182;  
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVVLGALALAGLGPAPAEPPQSGQCVHEHDCFALYPGPATFLNASQICDGLRHLM 60  
|||||  
Db 1 MLGVVLGALALAGLGPAPAEPPQSGQCVHEHDCFALYPGPATFLNASQICDGLRHLM 60  
QY 61 TVRSSVAADVISLLNDGGVGRRLRIGLQLPPCGDPRKRLGRLGFQWVTGDNNTSYS 120  
|||||  
Db 61 TVRSSVAADVISLLNDGGVGRRLRIGLQLPPCGDPRKRLGRLGFQWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGLPCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAY 180  
|||||  
Db 121 RWARLDLNGAPLCGLPCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAY 180  
QY 181 EPGAAAASVITYTTPFAARGADPQALPVGSSAAVAPLGLQLMCTAPPVAGQHWAREAP 240  
|||||  
Db 181 EPGAAAASVITYTTPFAARGADPQALPVGSSAAVAPLGLQLMCTAPPVAGQHWAREAP 240

QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300  
|||||  
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300  
QY 301 DOPGYSYSCMCTGYRLAADOHRCEVDVDDCILEPSPQPCQRCVNTQGGFECGCHYPNYDLVDG 360  
|||||  
Db 301 DOPGYSYSCMCTGYRLAADOHRCEVDVDDCILEPSPQPCQRCVNTQGGFECGCHYPNYDLVDG 360  
QY 361 ECVEPVDFCFRANCEYQCOPLNQTSLYCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
|||||  
Db 361 ECVEPVDFCFRANCEYQCOPLNQTSLYCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
QY 421 TQASCEPEGYILDDGFICTDDECENGFCGSGVCHNLPGTTECICGPDPSALVRH 475  
|||||  
Db 421 TQASCEPEGYILDDGFICTDDECENGFCGSGVCHNLPGTTECICGPDPSALVRH 475

RESULT 15  
US-08-587-389-2  
; Sequence 2, Application US/08587389  
; Patent No. 5695964  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI  
; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIFF & BERRIDGE  
; STREET: P.O. BOX 19928  
; CITY: ALEXANDRIA  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/587,389  
; FILING DATE: 17-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/307,444  
; FILING DATE: 19-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OLIFF, JAMES A.  
; REGISTRATION NUMBER: 27,075  
; REFERENCE/DOCKET NUMBER: JAO 27706  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6400  
; TELEFAX: (703) 836-2787  
; TELEX: 90-1799 PTO ALEX  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-587-389-2

Query Match 92.1%; Score 2686; DB 1; Length 475;  
Best Local Similarity 99.8%; Pred. No. 1.7e-182;  
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVVLGALALAGLGPAPAEPPQSGQCVHEHDCFALYPGPATFLNASQICDGLRHLM 60

1	MLGVLVJGALALAGLGGPAPAEPPQGGSGQVEHDCFALYPGPATFVNASQICDGLKGHLM	60
61	TVRSSVAADYISLLNGDGGVGRRRRLWIGLQPLPPGGCDPKRLGLRGFWTGDNNNTSYS	120
61	TVRSSVAADYISLLNGDGGVGRRRRLWIGLQPLPPGGCDPKRLGLRGFWTGDNNNTSYS	120
121	RWARLDLNGALPCLGPCLVAYSAAEATVPSEPIWEEQQCEVKADGFLCEFFHPFATCRPLAV	180
121	RWARLDLNGALPCLGPCLVAYSAAEATVPSEPIWEEQQCEVKADGFLCEFFHPFATCRPLAV	180
181	EPGAAAAAVSITTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAOVGHWAREAP	240
181	EPGAAAAAVSITTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAOVGHWAREAP	240
241	GAWDCSVENGSGEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP	300
241	GAWDCSVENGSGEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP	300
301	DQPSYSCMGETGYRLAADQHRCEDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360
301	DQPSYSCMGETGYRLAADQHRCEDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360
361	ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN	420
361	ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN	420
421	TQASCEPEGYILDDGFICTDIDECENGFCSCGVCHNLPGTETECIGPDSALVRH	475
421	TQASCEPEGYILDDGFICTDIDECENGFCSCGVCHNLPGTETECIGPDSALVRH	475

Search completed: December 16, 2002, 17:27:16  
Job time : 15.5 secs

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL; L39790; AAA74908.1; -	FT	DOMAIN	2484	2524	EGF-LIKE 42, CALCIUM-BINDING.
CC	EMBL; S69359; AAC60685.1; -	FT	DOMAIN	2525	2563	EGF-LIKE 43, CALCIUM-BINDING.
CC	HSSP; P35555; 1EMN.	FT	DOMAIN	2564	2606	EGF-LIKE 44, CALCIUM-BINDING.
DR	MGI:95490; Fbn2.	FT	DOMAIN	2607	2646	EGF-LIKE 45, CALCIUM-BINDING.
DR	InterPro; IPR000152; Asx_hydroxyl.	FT	DOMAIN	2647	2687	EGF-LIKE 46, CALCIUM-BINDING.
DR	InterPro; IPR001881; EGF-like.	FT	DOMAIN	2688	2727	EGF-LIKE 47, CALCIUM-BINDING.
DR	InterPro; IPR001438; EGF II.	FT	DOMAIN	115	124	BY SIMILARITY.
DR	InterPro; IPR002212; Fibril-assoc.	FT	DISULFID	119	130	BY SIMILARITY.
DR	Pfam; PF00008; EGF; 46.	FT	DISULFID	132	141	BY SIMILARITY.
DR	Pfam; PF00683; TB; 9.	FT	DISULFID	149	159	BY SIMILARITY.
DR	PRINTS; PRO0010; EGF-BLOOD.	FT	DISULFID	153	164	BY SIMILARITY.
DR	SMART; SM00179; EGF CA; 43.	FT	DISULFID	166	175	BY SIMILARITY.
DR	SMART; SM00001; EGF-like; 3.	FT	DISULFID	180	190	BY SIMILARITY.
DR	PROSITE; PS00010; ASX_HYDROXYL; 43.	FT	DISULFID	198	207	BY SIMILARITY.
DR	PROSITE; PS00022; EGF_1; 2.	FT	DISULFID	280	292	BY SIMILARITY.
DR	PROSITE; PS01186; EGF_2; 36.	FT	DISULFID	301	316	BY SIMILARITY.
DR	PROSITE; PS01187; EGF CA; 43.	FT	DISULFID	322	334	BY SIMILARITY.
DR	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;	FT	DISULFID	329	343	BY SIMILARITY.
KW	Repeat; Signal; Multigene family.	FT	DISULFID	345	358	BY SIMILARITY.
FT	POTENTIAL.	FT	DISULFID	491	503	BY SIMILARITY.
FT	CHAIN 1 28	FT	DISULFID	498	512	BY SIMILARITY.
FT	DOMAIN 29 2907	FT	DISULFID	514	526	BY SIMILARITY.
FT	EGF-LIKE 1.	FT	DISULFID	532	542	BY SIMILARITY.
FT	EGF-LIKE 2.	FT	DISULFID	537	551	BY SIMILARITY.
FT	EGF-LIKE 3.	FT	DISULFID	553	566	BY SIMILARITY.
FT	EGF-LIKE 4, CALCIUM-BINDING.	FT	DISULFID	572	584	BY SIMILARITY.
FT	EGF-LIKE 5, CALCIUM-BINDING.	FT	DISULFID	579	593	BY SIMILARITY.
FT	TGFBR 1.	FT	DISULFID	595	608	BY SIMILARITY.
FT	EGF-LIKE 6.	FT	DISULFID	614	625	BY SIMILARITY.
FT	EGF-LIKE 7, CALCIUM-BINDING.	FT	DISULFID	620	634	BY SIMILARITY.
FT	EGF-LIKE 8, CALCIUM-BINDING.	FT	DISULFID	636	649	BY SIMILARITY.
FT	EGF-LIKE 9, CALCIUM-BINDING.	FT	DISULFID	655	666	BY SIMILARITY.
FT	EGF-LIKE 10, CALCIUM-BINDING.	FT	DISULFID	661	675	BY SIMILARITY.
FT	TGFBR 2.	FT	DISULFID	677	690	BY SIMILARITY.
FT	EGF-LIKE 11, CALCIUM-BINDING.	FT	DISULFID	765	777	BY SIMILARITY.
FT	EGF-LIKE 12, CALCIUM-BINDING.	FT	DISULFID	772	786	BY SIMILARITY.
FT	EGF-LIKE 13, CALCIUM-BINDING.	FT	DISULFID	788	801	BY SIMILARITY.
FT	EGF-LIKE 14, CALCIUM-BINDING.	FT	DISULFID	807	819	BY SIMILARITY.
FT	TGFBR 3.	FT	DISULFID	814	828	BY SIMILARITY.
FT	EGF-LIKE 15, CALCIUM-BINDING.	FT	DISULFID	830	843	BY SIMILARITY.
FT	EGF-LIKE 16, CALCIUM-BINDING.	FT	DISULFID	849	859	BY SIMILARITY.
FT	EGF-LIKE 17, CALCIUM-BINDING.	FT	DISULFID	854	868	BY SIMILARITY.
FT	EGF-LIKE 18, CALCIUM-BINDING.	FT	DISULFID	870	883	BY SIMILARITY.
FT	EGF-LIKE 19, CALCIUM-BINDING.	FT	DISULFID	959	973	BY SIMILARITY.
FT	EGF-LIKE 20, CALCIUM-BINDING.	FT	DISULFID	975	988	BY SIMILARITY.
FT	EGF-LIKE 21, CALCIUM-BINDING.	FT	DISULFID	975	988	BY SIMILARITY.
FT	EGF-LIKE 22, CALCIUM-BINDING.	FT	DISULFID	1070	1082	BY SIMILARITY.
FT	EGF-LIKE 23, CALCIUM-BINDING.	FT	DISULFID	1077	1091	BY SIMILARITY.
FT	EGF-LIKE 24, CALCIUM-BINDING.	FT	DISULFID	1093	1106	BY SIMILARITY.
FT	EGF-LIKE 25, CALCIUM-BINDING.	FT	DISULFID	1112	1124	BY SIMILARITY.
FT	TGFBR 4.	FT	DISULFID	1119	1133	BY SIMILARITY.
FT	EGF-LIKE 26, CALCIUM-BINDING.	FT	DISULFID	1135	1149	BY SIMILARITY.
FT	EGF-LIKE 27, CALCIUM-BINDING.	FT	DISULFID	1155	1167	BY SIMILARITY.
FT	EGF-LIKE 28, CALCIUM-BINDING.	FT	DISULFID	1162	1176	BY SIMILARITY.
FT	TGFBR 5.	FT	DISULFID	1178	1191	BY SIMILARITY.
FT	EGF-LIKE 29, CALCIUM-BINDING.	FT	DISULFID	1197	1209	BY SIMILARITY.
FT	EGF-LIKE 30, CALCIUM-BINDING.	FT	DISULFID	1204	1218	BY SIMILARITY.
FT	EGF-LIKE 31, CALCIUM-BINDING.	FT	DISULFID	1220	1233	BY SIMILARITY.
FT	EGF-LIKE 32, CALCIUM-BINDING.	FT	DISULFID	1239	1250	BY SIMILARITY.
FT	EGF-LIKE 33, CALCIUM-BINDING.	FT	DISULFID	1246	1259	BY SIMILARITY.
FT	EGF-LIKE 34, CALCIUM-BINDING.	FT	DISULFID	1261	1274	BY SIMILARITY.
FT	EGF-LIKE 35, CALCIUM-BINDING.	FT	DISULFID	1280	1292	BY SIMILARITY.
FT	TGFBR 6.	FT	DISULFID	1287	1301	BY SIMILARITY.
FT	EGF-LIKE 36, CALCIUM-BINDING.	FT	DISULFID	1303	1316	BY SIMILARITY.
FT	EGF-LIKE 37, CALCIUM-BINDING.	FT	DISULFID	1322	1334	BY SIMILARITY.
FT	EGF-LIKE 38, CALCIUM-BINDING.	FT	DISULFID	1329	1343	BY SIMILARITY.
FT	EGF-LIKE 39, CALCIUM-BINDING.	FT	DISULFID	1345	1358	BY SIMILARITY.
FT	EGF-LIKE 40, CALCIUM-BINDING.	FT	DISULFID	1364	1377	BY SIMILARITY.
FT	TGFBR 7.	FT	DISULFID	1371	1386	BY SIMILARITY.
FT	EGF-LIKE 41, CALCIUM-BINDING.	FT	DISULFID	1388	1399	BY SIMILARITY.
FT	2442	FT	DISULFID	1405	1418	BY SIMILARITY.



FT DISULFID 814 827 BY SIMILARITY.  
FT DISULFID 821 836 BY SIMILARITY.  
FT DISULFID 843 856 BY SIMILARITY.  
FT DISULFID 862 875 BY SIMILARITY.  
FT DISULFID 869 884 BY SIMILARITY.  
FT DISULFID 886 899 BY SIMILARITY.  
FT DISULFID 905 917 BY SIMILARITY.  
FT DISULFID 913 926 BY SIMILARITY.  
FT DISULFID 928 941 BY SIMILARITY.  
FT DISULFID 947 956 BY SIMILARITY.  
FT DISULFID 952 965 BY SIMILARITY.  
FT DISULFID 967 980 BY SIMILARITY.  
FT DISULFID 986 998 BY SIMILARITY.  
FT DISULFID 994 1007 BY SIMILARITY.  
FT DISULFID 1009 1023 BY SIMILARITY.  
FT DISULFID 1029 1042 BY SIMILARITY.  
FT DISULFID 1036 1051 BY SIMILARITY.  
FT DISULFID 1056 1068 BY SIMILARITY.  
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1184 AA; 126543 MW; CA48490A55F98C5D CRC64;  
  
Query Match 12.68; Score 367.5; DB 1; Length 1184;  
Best Local Similarity 30.08; Pred. No. 4.7e-17;  
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;  
  
QY 154 EEOQCEVKAD-GFICEFHPATC-----RPLAV-----EPGAAAVSITYGTPFAARG 201  
Db 531 EQQCESNPNGPCN-HVMSLSCGEPLIVPEVRPPPEAPAPRRVS-----EAE 582  
  
QY 202 ADFQALPVGSSNAV---APLGLQMLCTAPPGAVGQHWAREAPGAWDCSVENGCGEHACNA 258  
Db 583 AGREALSLGTEAELPNSLPGDDQDCLELLPGEL-----COHLICIN 622  
  
QY 259 IPGAPRCQCPAGALQADGRSC-----TASAT-----QS 287  
Db 623 TVGSHCACFPFGLSLODDGRTRCPGHPQPEAPQEPALAKSEFSQVASTNTPLPLPQNP 682  
  
QY 288 CND--LCEHFCVNPDPQGVSCMCEGYRLAADOHRCEDVDDCILEPSPCP--QRCVNT 343  
Db 683 CKNGPCQKVC---STVGSAIGSCFFGYAIMADGVSCEDINECVTLHTCSRGEHCVNT 739  
  
QY 344 QGGFECH---CYPNYDLVDGCEVEPYDPGFRANCEYQCOPLNOTSILCV-----389  
Db 740 LGSFHCYKALTCEPGYALKDGC-EDVDEC--AMGTHQCP-----GFLCQNTKGSFYCQA 792  
  
QY 390 ---CAEGFAPIPH-----EPHRCQMFNCNQTACPADCPNTQASCEQPEGY-I 432  
Db 793 RQRCMDGFLQDPGNCVDINECTSLSEPCRPFGFSCINTVGSYTCQNPPLI---CARGYHA 849  
  
QY 433 LDGFICTDIDECENGFCSG---VCHNLPGTFECIC 466  
Db 850 SDDGAKVDVNEGTGVHRCGEGQVCHNLPGSVRCDC 886  
  
RESULT 9  
FBN1\_BOVIN STANDARD; PRT; 2871 AA.  
AC FBN1\_BOVIN  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrillin 1 precursor (MP340).  
GN FBN1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;

RA MEDLINE-95137597; PubMed-7835900;  
RA Tilstra D.J., Potter K.A., Byers P.H.;  
RT "Sequence of the coding region of the bovine fibrillin cDNA and  
RT localization to bovine chromosome 10.";  
RL Genomics 23:480-485(1994).  
RN [2]  
RP PARTIAL SEQUENCE.  
RX MEDLINE-96132851; PubMed-8557636;  
RA Gibson M.A., Hatzinikolas G., Kumaratilake J.S., Sandberg L.B.,  
RA Nicholl J.K., Sutherland G.R., Cleary E.G.;  
RT "Further characterization of proteins associated with elastic fiber  
RT microfibrils including the molecular cloning of MAGP-2 (MP25).";  
J. Biol. Chem. 271:1096-1103(1996).  
CC -|- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
CC -|- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER  
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE  
CC MICROFIBRILS.  
CC -|- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.  
CC -|- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L28748; AAA74122.1; -;  
DR HSP; P35555; LAPJ.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF-Ca.  
DR InterPro; IPR001438; EGF-II.  
DR InterPro; IPR002212; Fibrill-assoc.  
DR Pfam; PF00008; EGF; 46.  
DR Pfam; PF00683; TB; 9.  
DR PRINTS; PR001010; EGFBL00D.  
DR SMART; SM00179; EGF\_CA; 42.  
DR SMART; SM00001; EGF\_Like; 4.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 43.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 38.  
DR PROSITE; PS01187; EGF\_CA; 43.  
DR Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
DR Repeat; Signal; Multigene family.  
FT SIGNAL 1 27 POTENTIAL.  
FT CHAIN 28 2871 FIBRILLIN 1.  
FT DOMAIN 81 112 EGF-LIKE 1.  
FT DOMAIN 115 146 EGF-LIKE 2.  
FT DOMAIN 147 178 EGF-LIKE 3.  
FT DOMAIN 246 287 EGF-LIKE 4.  
FT DOMAIN 288 329 CALCIUM-BINDING.  
FT REPEAT 330 390 EGF-LIKE 5, CALCIUM-BINDING.  
FT REPEAT 392 446 TGFBP 1.  
FT DOMAIN 449 489 PRO-RICH.  
FT DOMAIN 490 529 EGF-LIKE 6.  
FT DOMAIN 530 571 EGF-LIKE 7.  
FT DOMAIN 572 612 EGF-LIKE 8, CALCIUM-BINDING.  
FT DOMAIN 613 653 EGF-LIKE 9, CALCIUM-BINDING.  
FT REPEAT 654 722 EGF-LIKE 10, CALCIUM-BINDING.  
FT DOMAIN 723 764 TGFBP 2.  
FT DOMAIN 765 806 EGF-LIKE 11, CALCIUM-BINDING.  
FT DOMAIN 807 846 EGF-LIKE 12, CALCIUM-BINDING.  
FT DOMAIN 910 951 EGF-LIKE 13, CALCIUM-BINDING.  
FT REPEAT 952 1027 TGFBP 3.  
FT DOMAIN 1028 1069 EGF-LIKE 14, CALCIUM-BINDING.  
FT DOMAIN 1070 1112 EGF-LIKE 15, CALCIUM-BINDING.  
FT DOMAIN 1113 1154 EGF-LIKE 16, CALCIUM-BINDING.  
FT DOMAIN 1155 1196 EGF-LIKE 17, CALCIUM-BINDING.  
FT DOMAIN 1197 EGF-LIKE 18, CALCIUM-BINDING.  
FT DOMAIN 1237 EGF-LIKE 19, CALCIUM-BINDING.

```
FT DOMAIN 1238 EGF-LIKE 20, CALCIUM-BINDING. 1279
FT DOMAIN 1280 EGF-LIKE 21, CALCIUM-BINDING. 1321
FT DOMAIN 1322 EGF-LIKE 22, CALCIUM-BINDING. 1362
FT DOMAIN 1363 EGF-LIKE 23, CALCIUM-BINDING. 1403
FT DOMAIN 1404 EGF-LIKE 24, CALCIUM-BINDING. 1445
FT DOMAIN 1446 EGF-LIKE 25, CALCIUM-BINDING. 1486
FT DOMAIN 1487 EGF-LIKE 26, CALCIUM-BINDING. 1527
FT REPEAT 1528 TGFBP 4. 1605
FT DOMAIN 1606 EGF-LIKE 27, CALCIUM-BINDING. 1647
FT DOMAIN 1648 EGF-LIKE 28, CALCIUM-BINDING. 1688
FT REPEAT 1689 TGFBP 5. 1765
FT DOMAIN 1766 EGF-LIKE 29, CALCIUM-BINDING. 1807
FT DOMAIN 1808 EGF-LIKE 30, CALCIUM-BINDING. 1848
FT DOMAIN 1849 EGF-LIKE 31, CALCIUM-BINDING. 1890
FT DOMAIN 1891 EGF-LIKE 32, CALCIUM-BINDING. 1929
FT DOMAIN 1930 EGF-LIKE 33, CALCIUM-BINDING. 1972
FT DOMAIN 1973 EGF-LIKE 34, CALCIUM-BINDING. 2012
FT DOMAIN 2013 EGF-LIKE 35, CALCIUM-BINDING. 2054
FT REPEAT 2055 TGFBP 6. 2126
FT DOMAIN 2127 EGF-LIKE 36, CALCIUM-BINDING. 2165
FT DOMAIN 2166 EGF-LIKE 37, CALCIUM-BINDING. 2205
FT DOMAIN 2206 EGF-LIKE 38, CALCIUM-BINDING. 2246
FT DOMAIN 2247 EGF-LIKE 39, CALCIUM-BINDING. 2290
FT DOMAIN 2291 EGF-LIKE 40, CALCIUM-BINDING. 2332
FT REPEAT 2333 TGFBP 7. 2401
FT DOMAIN 2402 EGF-LIKE 41, CALCIUM-BINDING. 2443
FT DOMAIN 2444 EGF-LIKE 42, CALCIUM-BINDING. 2484
FT DOMAIN 2485 EGF-LIKE 43, CALCIUM-BINDING. 2523
FT DOMAIN 2524 EGF-LIKE 44, CALCIUM-BINDING. 2566
FT DOMAIN 2567 EGF-LIKE 45, CALCIUM-BINDING. 2606
FT DOMAIN 2607 EGF-LIKE 46, CALCIUM-BINDING. 2647
FT DOMAIN 2648 EGF-LIKE 47, CALCIUM-BINDING. 2687
FT DISULFID 85 BY SIMILARITY. 94
FT DISULFID 89 BY SIMILARITY. 100
FT DISULFID 102 BY SIMILARITY. 111
FT DISULFID 119 BY SIMILARITY. 129
FT DISULFID 123 BY SIMILARITY. 134
FT DISULFID 136 BY SIMILARITY. 145
FT DISULFID 150 BY SIMILARITY. 160
FT DISULFID 154 BY SIMILARITY. 166
FT DISULFID 168 BY SIMILARITY. 177
FT DISULFID 250 BY SIMILARITY. 262
FT DISULFID 257 BY SIMILARITY. 271
FT DISULFID 273 BY SIMILARITY. 286
FT DISULFID 292 BY SIMILARITY. 304
FT DISULFID 299 BY SIMILARITY. 313
FT DISULFID 315 BY SIMILARITY. 328
FT DISULFID 453 BY SIMILARITY. 465
FT DISULFID 460 BY SIMILARITY. 474
FT DISULFID 476 BY SIMILARITY. 488
FT DISULFID 494 BY SIMILARITY. 504
FT DISULFID 499 BY SIMILARITY. 513
FT DISULFID 515 BY SIMILARITY. 528
FT DISULFID 534 BY SIMILARITY. 546
FT DISULFID 541 BY SIMILARITY. 555
FT DISULFID 557 BY SIMILARITY. 570
FT DISULFID 576 BY SIMILARITY. 587
FT DISULFID 582 BY SIMILARITY. 596
FT DISULFID 598 BY SIMILARITY. 611
FT DISULFID 617 BY SIMILARITY. 628
FT DISULFID 623 BY SIMILARITY. 637
FT DISULFID 639 BY SIMILARITY. 652
FT DISULFID 727 BY SIMILARITY. 739
FT DISULFID 734 BY SIMILARITY. 748
FT DISULFID 750 BY SIMILARITY. 763
FT DISULFID 769 BY SIMILARITY. 781
FT DISULFID 776 BY SIMILARITY. 790
FT DISULFID 792 BY SIMILARITY. 805
FT DISULFID 811 BY SIMILARITY. 821
FT DISULFID 816 BY SIMILARITY. 830
FT DISULFID 832 BY SIMILARITY. 845
FT DISULFID 914 BY SIMILARITY. 926
FT DISULFID 921 BY SIMILARITY. 935
```

```
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1166 1180 BY SIMILARITY.
FT DISULFID 1182 1195 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1431 1444 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.

Query Match 12.6%; Score 367; DB 1; Length 2871;
Best Local Similarity 31.5%; Pred. No. 1.2e-16;
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

QY 244 DCSVNGGCEHACNAIPGAPRCQCPAGAAIQADGRSCTASATSCNDLCEHFCVNP--- 300
DB 1200 ECSIMNGGCETFTCTNSGSEYECSCQPGFALMPDQRSC-----DIDE--CEDNENIC 1249
QY 301 -----DOPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPC-PQRCVNTGGFECHCYP 353
DB 1250 DGGQCTNIPGEYRCLCYDGFMASEDKMTKVDVNECDLNPNICLSGTGTCENTKGSFICHCDM 1309
QY 354 NYDLVDGE--CVRPVDPC--FRANCEYQCQPLNOT--SYLCVCAEGFA-----PIP 398
DB 1310 GYSGKKGKTGCTD-INCEIGAHNCDRHAVCTNTAGSFKSCSPGWICDGIKCTDIDEC 1368
QY 399 HEPHRCOMFCNQACPADCDPNTQAS--CECPGYILDDGFICTDIDEC-EN----- 447
DB 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEGY-IGDGTCTDIDECSENLCNGNQ 1419
QY 448 -----GGF-----CS-----GVCHNLPTGTEFCICGPDLSALV 473
DB 1420 CLNAPGGYRCEDMGFVPSADGKACEDIDECSLFNICVFGTCHNLPLGRCECEIGVELD 1479
QY 474 RHIG-----TDCDSK-VDGGDSGSGEPPTGSLTTPAVGLV 513
DB 1480 RSGGNCTDVNECLDPTTCISGNCVNTPGSYTCDCPPD-----FELNPRVGCV 1527

RESULT 10
FBNI_PIG
ID FBNI_PIG STANDARD; PRT; 2871 AA.
AC 09TV36;
```

DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Fibrillin 1 precursor.  
 GN FN1.  
 OS Sus scrofa (pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=99156858; PubMed=10036187;  
 RA Biery N.J.; Eldadah Z.A.; Moore C.S.; Stetten G.; Spencer F.;  
 RA Dietz H.C.;  
 RT \*Revised genomic organization of FN1 and significance for regulated  
 RL gene expression.";  
 RL Genomics 56:70-77(1999).  
 CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
 CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
 CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
 CC -1- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER  
 CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE  
 CC MICROFIBRILS.  
 CC -1- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF073800; AAD50328.1; .  
 DR HSSP; P35553; IAPJ.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001438; EGF\_II.  
 DR InterPro; IPR002212; Fibril-assoc.  
 DR Pfam; PF00008; EGF\_45.  
 DR Pfam; PF00683; TB; 9.  
 DR PRINTS; PR00010; EGFBL00D.  
 DR SMART; SM00179; EGF\_CA; 40.  
 DR SMART; SM00001; EGF\_like; 6.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 41.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_2.  
 DR PROSITE; PS01186; EGF\_2; 36.  
 DR PROSITE; PS01187; EGF\_CA; 41.  
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
 KW Repeat; Signal; Multigene family.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 2871 FIBRILLIN 1.  
 FT DOMAIN 81 112 EGF-LIKE 1.  
 FT DOMAIN 115 145 EGF-LIKE 2.  
 FT DOMAIN 147 178 EGF-LIKE 3.  
 FT DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.  
 FT DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.  
 FT REPEAT 330 390 TGFBP 1.  
 FT DOMAIN 392 445 PRO-RICH.  
 FT DOMAIN 449 489 EGF-LIKE 6.  
 FT DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.  
 FT DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.  
 FT DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.  
 FT DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.  
 FT REPEAT 654 722 TGFBP 2.  
 FT DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.  
 FT DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.  
 FT DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.  
 FT DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.  
 FT REPEAT 952 1027 TGFBP 3.  
 FT DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.  
 FT DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.  
 FT DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.  
 FT DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.  
 FT DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.  
 FT DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.  
 FT DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.  
 FT DOMAIN 1322 1363 EGF-LIKE 22, CALCIUM-BINDING.  
 FT DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.  
 FT DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.  
 FT DOMAIN 1446 1486 EGF-LIKE 25, CALCIUM-BINDING.  
 FT REPEAT 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.  
 FT REPEAT 1528 1605 TGFBP 4.  
 FT DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.  
 FT DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.  
 FT REPEAT 1689 1765 TGFBP 5.  
 FT DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.  
 FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.  
 FT DOMAIN 1849 1890 EGF-LIKE 31, CALCIUM-BINDING.  
 FT DOMAIN 1891 1929 EGF-LIKE 32, CALCIUM-BINDING.  
 FT DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.  
 FT DOMAIN 1973 2012 EGF-LIKE 34, CALCIUM-BINDING.  
 FT DOMAIN 2013 2054 EGF-LIKE 35, CALCIUM-BINDING.  
 FT REPEAT 2055 2126 TGFBP 6.  
 FT DOMAIN 2127 2165 EGF-LIKE 36, CALCIUM-BINDING.  
 FT DOMAIN 2166 2205 EGF-LIKE 37, CALCIUM-BINDING.  
 FT DOMAIN 2206 2246 EGF-LIKE 38, CALCIUM-BINDING.  
 FT DOMAIN 2247 2290 EGF-LIKE 39, CALCIUM-BINDING.  
 FT DOMAIN 2291 2332 EGF-LIKE 40, CALCIUM-BINDING.  
 FT REPEAT 2333 2401 TGFBP 7.  
 FT DOMAIN 2402 2443 EGF-LIKE 41, CALCIUM-BINDING.  
 FT DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.  
 FT DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.  
 FT DOMAIN 2524 2566 EGF-LIKE 44, CALCIUM-BINDING.  
 FT DOMAIN 2567 2606 EGF-LIKE 45, CALCIUM-BINDING.  
 FT DOMAIN 2607 2647 EGF-LIKE 46, CALCIUM-BINDING.  
 FT DOMAIN 2648 2687 EGF-LIKE 47, CALCIUM-BINDING.  
 FT DISULFID 85 94 BY SIMILARITY.  
 FT DISULFID 89 100 BY SIMILARITY.  
 FT DISULFID 102 111 BY SIMILARITY.  
 FT DISULFID 119 129 BY SIMILARITY.  
 FT DISULFID 123 134 BY SIMILARITY.  
 FT DISULFID 136 145 BY SIMILARITY.  
 FT DISULFID 150 160 BY SIMILARITY.  
 FT DISULFID 154 166 BY SIMILARITY.  
 FT DISULFID 168 177 BY SIMILARITY.  
 FT DISULFID 250 262 BY SIMILARITY.  
 FT DISULFID 257 271 BY SIMILARITY.  
 FT DISULFID 273 286 BY SIMILARITY.  
 FT DISULFID 292 304 BY SIMILARITY.  
 FT DISULFID 299 313 BY SIMILARITY.  
 FT DISULFID 315 328 BY SIMILARITY.  
 FT DISULFID 453 465 BY SIMILARITY.  
 FT DISULFID 460 474 BY SIMILARITY.  
 FT DISULFID 476 488 BY SIMILARITY.  
 FT DISULFID 494 504 BY SIMILARITY.  
 FT DISULFID 499 513 BY SIMILARITY.  
 FT DISULFID 515 528 BY SIMILARITY.  
 FT DISULFID 534 546 BY SIMILARITY.  
 FT DISULFID 541 555 BY SIMILARITY.  
 FT DISULFID 557 570 BY SIMILARITY.  
 FT DISULFID 576 587 BY SIMILARITY.  
 FT DISULFID 582 596 BY SIMILARITY.  
 FT DISULFID 598 611 BY SIMILARITY.  
 FT DISULFID 617 628 BY SIMILARITY.  
 FT DISULFID 623 637 BY SIMILARITY.  
 FT DISULFID 639 652 BY SIMILARITY.  
 FT DISULFID 727 739 BY SIMILARITY.  
 FT DISULFID 734 748 BY SIMILARITY.  
 FT DISULFID 750 763 BY SIMILARITY.  
 FT DISULFID 769 781 BY SIMILARITY.  
 FT DISULFID 776 790 BY SIMILARITY.  
 FT DISULFID 792 805 BY SIMILARITY.







QY 302 -----QPGSCMCTGYRLAADQHRQEDVDDCILEPSCP-QRCVNTQGGFECYCP 353  
DB 1294 DGGQCTNIGFGRCLYDGFMSMDKMTCDVNECDLNSNICMFGECKTGSFICHQCL 1353  
QY 354 NYDLVDGE--CYPEVDPC--FRANCYQCOPLN-QTSYLCVCAEGFAPTH-----EPHR 403  
DB 1354 GYSVKKGTTGCTD-VDECEIGAHCNDHASCNLNPGSFKSCREGW--IGNGIKCIDLDE 1410  
QY 404 COMFCNQTCAPDCDNTQAS--CECPGYLDGDFICTDIDE-----CENG----- 448  
DB 1411 CSNGTHQCSINAQC-VNTPGYSRCACSEGF-TGDGFTCSVDDECAENINLNCENGQCLNVP 1468  
QY 449 -----GF-----CSGVCHNLPGFECTCGPDSALVRHIG- 477  
DB 1469 GAYRCCEMGFTPASDSRSCQDIDECFQNICVSGTCNNLPGMFHCLICDDGYELDRGTGN 1528  
QY 478 -TDCD 481  
DB 1529 CTDID 1533  
PRT; 2871 AA.  
RESULT 12  
ID FBNI\_HUMAN STANDARD; PRT; 2871 AA.  
AC P35555;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrillin 1 precursor.  
GN FBNI OR FBNI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX MEDLINE=93372860; PubMed=8364578;  
RA Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,  
Pangillinan T., Bonadio J.;  
RT "Genomic organization of the sequence coding for fibrillin, the  
defective gene product in Marfan syndrome.";  
RL Hum. Mol. Genet. 2:961-968(1993).  
RN [2]  
RP SEQUENCE OF 1-932 FROM N.A.  
RC TISSUE=Placenta, and Fibroblast;  
RX MEDLINE=94010947; PubMed=7691719;  
RA Corson G.M., Chalberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;  
RT "Fibrillin binds calcium and is coded by cDNAs that reveal a  
multidomain structure and alternatively spliced exons at the 5'  
end.";  
RL Genomics 17:476-484(1993).  
RN [3]  
RP SEQUENCE OF 899-2871 FROM N.A.  
RX MEDLINE=91304568; PubMed=1852207;  
RA Maslen C.L., Corson G.M., Maddox B.K., Glanville R.W., Sakai L.Y.;  
RT "Partial sequence of a candidate gene for the Marfan syndrome.";  
RL Nature 352:334-337(1991).  
RN [4]  
RP SEQUENCE OF 813-1313 FROM N.A.  
RX MEDLINE=91304567; PubMed=1852206;  
RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,  
Tsipouras P., Ramirez F., Hollister D.W.;  
RT "Linkage of Marfan syndrome and a phenotypically related disorder to  
two different fibrillin genes.";  
RL Nature 352:330-334(1991).  
RN [5]  
RP CHARACTERIZATION.  
RX MEDLINE=91317849; PubMed=1860873;  
RA Sakai L.Y., Keene D.R., Glanville R.W., Bachinger H.P.;  
RT "Purification and partial characterization of fibrillin, a cysteine-  
rich structural component of connective tissue microfibrils.";

RL J. Biol. Chem. 266:14763-14770(1991).  
RN [6]  
RP STRUCTURE BY NMR OF 2054-2125.  
RX MEDLINE=98031893; PubMed=9362480;  
RA Yuan X., Downing A.K., Knott V., Handford P.A.;  
RT "Solution structure of the transforming growth factor beta-binding  
protein-like module, a domain associated with matrix fibrils.";  
RL EMBO J. 16:6659-6666(1997).  
RN [7]  
RP STRUCTURE BY NMR OF 2124-2205.  
RX MEDLINE=96144829; PubMed=8568869;  
RA Knott V., Downing A.K., Cardy C.M., Handford P.A.;  
RT "Calcium binding properties of an epidermal growth factor-like domain  
pair from human fibrillin-1.";  
RL J. Mol. Biol. 255:22-27(1996).  
RN [8]  
RP STRUCTURE BY NMR OF 2124-2205.  
RX MEDLINE=96222301; PubMed=8653794;  
RA Downing A.K., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,  
Handford P.A.;  
RT "Solution structure of a pair of calcium-binding epidermal growth  
factor-like domains: implications for the Marfan syndrome and other  
genetic disorders.";  
RL Cell 85:597-603(1996).  
RN [9]  
RP REVIEW ON MFS VARIANTS.  
RX MEDLINE=96174615; PubMed=8594563;  
RA Colod G., Beroud C., Soussi T., Junien C., Boileau C.;  
RT "Software and database for the analysis of mutations in the human  
FBNI gene.";  
RL Nucleic Acids Res. 24:137-141(1996).  
RN [10]  
RP REVIEW ON MFS VARIANTS.  
RX MEDLINE=97169383; PubMed=9016526;  
RA Colod-Beroud G., Beroud C., Ades L., Black C., Boxer M., Brock D.J.,  
Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,  
Richards R.I., Wang W., Junien C., Boileau C.;  
RT "Marfan Database (second edition): software and database for the  
analysis of mutations in the human FBNI gene.";  
RL Nucleic Acids Res. 25:147-150(1997).  
RN [11]  
RP REVIEW ON VARIANTS.  
RX MEDLINE=98062175; PubMed=9401003;  
RA Hayward C., Brock D.J.H.;  
RT "Fibrillin-1 mutations in Marfan syndrome and other type-1  
fibrillinopathies.";  
RL Hum. Mutat. 10:415-423(1997).  
RN [12]  
RP VARIANT MFS PRO-1137.  
RX MEDLINE=91304569; PubMed=1852208;  
RA Dietz H.C., Cutting G.R., Peyeritz R.E., Maslen C.L., Sakai L.Y.,  
Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,  
Curristin S.M., Stetten G., Meyers D.A., Francomano C.A.;  
RT "Marfan syndrome caused by a recurrent de novo missense mutation in  
the fibrillin gene.";  
RL Nature 352:337-339(1991).  
RN [13]  
RP VARIANTS MFS SER-1249; ARG-1663; SER-2221 AND SER-2307.  
RX MEDLINE=93250834; PubMed=1301946;  
RA Dietz H.C., Saraiva J.M., Peyeritz R.E., Cutting G.R., Francomano C.A.;  
RT "Clustering of fibrillin (FBNI) missense mutations in Marfan syndrome  
patients at cysteine residues in EGF-like domains.";  
RL Hum. Mutat. 1:366-374(1992).  
RN [14]  
RP VARIANT MFS SER-2307.  
RX MEDLINE=92235290; PubMed=1569206;  
RA Dietz H.C., Peyeritz R.E., Puffenberger E.G., Kendzior R.J. Jr.,  
Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;  
RT "Marfan phenotype variability in a family segregating a missense  
mutation in the epidermal growth factor-like motif of the fibrillin  
gene.";  
RL J. Clin. Invest. 89:1674-1680(1992).  
RN [15]

RP VARIANTS MFS ILE-548 AND ALA-723.  
 RX MEDLINE-94010946; PubMed-8406497;  
 RA Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C.,  
 RA Pyeritz R.E., Francomano C.A.;  
 RT "Four novel FBNI mutations: significance for mutant transcript level  
 RT and EGF-like domain calcium binding in the pathogenesis of Marfan  
 RT syndrome.";  
 RL Genomics 17:468-475(1993).  
 RN [16]  
 RN VARIANTS MFS SER-2144.  
 RX MEDLINE-93378402; PubMed-8504310;  
 RA Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;  
 RT "A novel fibrillin mutation in the Marfan syndrome which could  
 RT disrupt calcium binding of the epidermal growth factor-like module.";  
 RL Hum. Mol. Genet. 2:475-477(1993).  
 RN [17]  
 RN VARIANTS MFS R-862; Y-1117; P-1137 AND F-1589, AND VARIANT A-1148.  
 RX MEDLINE-94108431; PubMed-8281141;  
 RA Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Ganser C.,  
 RA Berg M.A., Miller D.C., Francke U.;  
 RT "Mutation screening of complete fibrillin-1 coding sequence: report  
 RT of five new mutations, including two in 8-cysteine domains.";  
 RL Hum. Mol. Genet. 2:1813-1821(1993).  
 RN [18]  
 RN VARIANTS MFS GLY-217 AND ARG-2627.  
 RX MEDLINE-95067970; PubMed-7977366;  
 RA Karttunen L., Raghunath M., Loenqvist L., Peltonen L.;  
 RT "A compound-heterozygous Marfan patient: two defective fibrillin  
 RT alleles result in a lethal phenotype.";  
 RL Am. J. Hum. Genet. 55:1083-1091(1994).  
 RN [19]  
 RN VARIANT EL LYS-2447.  
 RX MEDLINE-94245249; PubMed-8188302;  
 RA Lonqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,  
 RA Peltonen L.;  
 RT "A novel mutation of the fibrillin gene causing ectopia lentis.";  
 RL Genomics 19:573-576(1994).  
 RN [20]  
 RN VARIANT MFS CYS-627.  
 RX MEDLINE-94272487; PubMed-8004112;  
 RA Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;  
 RT "Two novel mutations and a neutral polymorphism in EGF-like domains  
 RT of the fibrillin gene (FBNI): SSCP screening of exons 15-21 in Marfan  
 RT syndrome patients.";  
 RL Hum. Mol. Genet. 3:373-375(1994).  
 RN [21]  
 RN VARIANT MFS CYS-122.  
 RX MEDLINE-94314977; PubMed-8040326;  
 RA Stahl-Hallengren C., Ukkonen T., Kainulainen K., Kristofersson U.,  
 RA Saxne I., Tornqvist K., Peltonen L.;  
 RT "An extra cysteine in one of the non-calcium-binding epidermal growth  
 RT factor-like motifs of the FBNI polypeptide is connected to a novel  
 RT variant of Marfan syndrome.";  
 RL J. Clin. Invest. 94:709-713(1994).  
 RN [22]  
 RN VARIANT MFS TRP-1223.  
 RX MEDLINE-94351682; PubMed-8071963;  
 RA Hewett D.R., Lynch J.R., Child A., Sykes B.C.;  
 RT "A new missense mutation of fibrillin in a patient with Marfan  
 RT syndrome.";  
 RL J. Med. Genet. 31:338-339(1994).  
 RN [23]  
 RN VARIANT MFS HIS-1170.  
 RX MEDLINE-95174777; PubMed-7870075;  
 RA Hayward C., Porteous M.E.M., Brock D.J.H.;  
 RT "A novel mutation in the fibrillin gene (FBNI) in familial  
 RT arachnodactyly.";  
 RL Mol. Cell. Probes 8:325-327(1994).  
 RN [24]  
 RN VARIANTS MFS G-217; N-1023; R-1074; Y-1242; R-1513; E-2127; W-2151;  
 RP K-2447 AND R-2511.  
 RX MEDLINE-94184368; PubMed-8136837;  
 RA Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.;

RT "Mutations in the fibrillin gene responsible for dominant ectopia  
 RT lentis and neonatal Marfan syndrome.";  
 RL Nat. Genet. 6:64-69(1994).  
 RN [25]  
 RP VARIANT SER-1127.  
 Query Match 12.4%; Score 363; DB 1; Length 2871;  
 Best Local Similarity 30.8%; Pred. NO. 2.1e-16;  
 Matches 111; Conservative 22; Mismatches 105; Indels 122; Gaps 20;  
 QY 244 DCSVNGGCEHACNAIPGAPRCQCPAGALADGRSCSTASATQSCNDLCEHFCVNP--- 300  
 DB 1200 ECSIWNGGCEFTCTNSSESYECSCQGFALMPDQRSCT-----DIDE--CEDNPIC 1249  
 QY 301 -----DOPGSYSMCETGYRLAADQHRCEVDVDDCILPSPC--PORCVNTOGGFCHCYP 353  
 DB 1250 DGGQCTNIPGEYRCLYCYGDFMASEDMKTCVDVNECDLNPICSLCTCKTSGSFICHCDM 1309  
 QY 354 NYDLVDGE--CVPEDVPC--FRANCEYQCPUNQT-SYLCVCAEFA-----PIP 398  
 DB 1310 GYSGRKGKGTCTD-INCEIGAHCNGKHAVCTNTAGSFKSCSPGWIGDKICTDLDECS 1368  
 QY 399 HEPHRCQMFECNQACPDADPNTQAS--CECPGYLDDGFICTDIDEC-EN----- 447  
 DB 1369 NGTHMCSQH-----ADC-KNTMGSYRLCKEGY-TGDFCTDLDECSNLNLCGNGQ 1419  
 QY 448 -----GGF-----CS-----GVCHNLPGTFFECICGPDLSALV 473  
 DB 1420 CLNAPGGYRCEDMGFVPSADGKACEDDECSLPNICVFGTCHNLPGLFRCECEIGYELD 1479  
 QY 474 RHIG-----TDCDSGKVGDSGSGSEPPSPTPGS-----TLTPPAVGLV 513  
 DB 1480 RSGGNCITDNECLDPTTCISGCVN-----TPGSYICDCPPDFELNPTRVGCV 1527  
 RESULT 13  
 FBNI\_MOUSE  
 ID FBNI\_MOUSE STANDARD; PRT; 2871 AA.  
 AC Q61554; Q60826;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Fibrillin 1 precursor.  
 GN FBNI OR FBNI-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID-10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-95130561; PubMed-7829516;  
 RA Yin W., Germiller J., Sanguinetti C., Smiley E., Panglilan T.,  
 RA Pereira L., Ramirez F., Bonadio J.;  
 RT "Primary structure and developmental expression of Fbn-1, the mouse  
 RT fibrillin gene.";  
 RL J. Biol. Chem. 270:1798-1806(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CD-1; TISSUE-Kidney;  
 RA Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;  
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
 CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
 CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
 CC -!- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER  
 CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE  
 CC MICROFIBRILS (BY SIMILARITY).  
 CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its









FT	DOMAIN	511	543	ANAPHYLATOXIN-LIKE 3.
FT	DOMAIN	594	635	EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	669	708	EGF-LIKE 2.
FT	DOMAIN	709	755	EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	756	800	EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	801	846	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	847	894	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	895	937	EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	938	979	EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	980	1018	EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1019	1061	EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1062	1106	EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1111	1221	DOMAIN III.
FT	SITE	421	423	CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID	435	462	BY SIMILARITY.
FT	DISULFID	436	469	BY SIMILARITY.
FT	DISULFID	449	470	BY SIMILARITY.
FT	DISULFID	479	508	BY SIMILARITY.
FT	DISULFID	492	509	BY SIMILARITY.
FT	DISULFID	511	535	BY SIMILARITY.
FT	DISULFID	512	542	BY SIMILARITY.
FT	DISULFID	525	543	BY SIMILARITY.
FT	DISULFID	598	610	BY SIMILARITY.
FT	DISULFID	606	619	BY SIMILARITY.
FT	DISULFID	621	634	BY SIMILARITY.
FT	DISULFID	673	683	BY SIMILARITY.
FT	DISULFID	679	692	BY SIMILARITY.
FT	DISULFID	694	707	BY SIMILARITY.
FT	DISULFID	713	726	BY SIMILARITY.
FT	DISULFID	720	735	BY SIMILARITY.
FT	DISULFID	742	754	BY SIMILARITY.
FT	DISULFID	805	818	BY SIMILARITY.
FT	DISULFID	812	827	BY SIMILARITY.
FT	DISULFID	833	845	BY SIMILARITY.
FT	DISULFID	899	912	BY SIMILARITY.
FT	DISULFID	906	921	BY SIMILARITY.
FT	DISULFID	923	936	BY SIMILARITY.
FT	DISULFID	942	954	BY SIMILARITY.
FT	DISULFID	950	963	BY SIMILARITY.
FT	DISULFID	965	978	BY SIMILARITY.
FT	DISULFID	984	993	BY SIMILARITY.
FT	DISULFID	989	1002	BY SIMILARITY.
FT	DISULFID	1004	1017	BY SIMILARITY.
FT	DISULFID	1023	1035	BY SIMILARITY.
FT	DISULFID	1031	1044	BY SIMILARITY.
FT	DISULFID	1046	1060	BY SIMILARITY.
FT	DISULFID	1066	1079	BY SIMILARITY.
FT	DISULFID	1073	1088	BY SIMILARITY.
FT	DISULFID	1093	1105	BY SIMILARITY.
FT	CARBOHYD	179	179	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	497	497	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	737	737	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1072	1072	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARSPLIC	709	755	MISSING (IN ISOFORM 2).
FT	CONFLICT	140	159	HSGRKYAGGHVHVLSSCRAC -> TVAVSICWVYRPLLIPL
FT	CONFLICT	348	348	GF (IN REF. 2).
FT	CONFLICT	507	507	S -> L (IN REF. 2).
FT	CONFLICT	1102	1102	Q -> QQ (IN REF. 2).
FT	CONFLICT	1221	1221	Q -> E (IN REF. 2).
FT	CONFLICT	131818	131818	MM; 87B92A10A9FDc45F CRC64
FT	CONFLICT	1221	1221	AA; 131818
FT	CONFLICT	140	159	GF (IN REF. 2).
FT	CONFLICT	348	348	S -> L (IN REF. 2).
FT	CONFLICT	507	507	Q -> QQ (IN REF. 2).
FT	CONFLICT	1102	1102	Q -> E (IN REF. 2).
FT	CONFLICT	131818	131818	MM; 87B92A10A9FDc45F CRC64
FT	CONFLICT	1221	1221	AA; 131818
FT	CONFLICT	140	159	GF (IN REF. 2).
FT	CONFLICT	348	348	S -> L (IN REF. 2).
FT	CONFLICT	507	507	Q -> QQ (IN REF. 2).
FT	CONFLICT	1102	1102	Q -> E (IN REF. 2).
FT	CONFLICT	131818	131818	MM; 87B92A10A9FDc45F CRC64
FT	CONFLICT	1221	1221	AA; 131818
FT	CONFLICT	140	159	GF (IN REF. 2).
FT	CONFLICT	348	348	S -> L (IN REF. 2).
FT	CONFLICT	507	507	Q -> QQ (IN REF. 2).
FT	CONFLICT	1102	1102	Q -> E (IN REF. 2).
FT	CONFLICT	131818	131818	MM; 87B92A10A9FDc45F CRC64
FT	CONFLICT	1221	1221	AA; 131818
FT	CONFLICT	140	159	GF (IN REF. 2).
FT	CONFLICT	348	348	S -> L (IN REF. 2).
FT	CONFLICT	507	507	

Qy	230	AVQGHWARE--APGAWDCSV-----ENGCGEHCACNAIPCAPRCQCPAGAA	272
Db	640	APQLDTRASAPRSESQAQVSPNTIPLVPVPNTCKDNGPCRCRVVGGDTAMCSCFFPGYA	699
Qy	273	LOADGRSC-----TASATQSCNDLCBHFVPNDQPGSYSC-----MCETGYRLAADQR	322
Db	700	IMADVSCEDQDECLMTHDCS--WKQFCV----NTLGSFVCVNHTVLCAEYILNA--HRK	753
Qy	323	CEDVDDCILPESPC--PQRCVNTQGFBECH-----CYPNTDLDVGECEVPVDPFCFRA--NC	374
Db	754	CVDINECVTLDLHTCTRAEHCVNTPGSGFQCYKALTCEPFGYVLTGDGCTD--VDECVTGTIHC	812
Qy	375	E--YQOCPLNOTSYLCV-----CAEGFAPIPH-----EPHRQCMFCNOTACPA	415
Db	813	QAGFSQCN--YGSFYCQARQCMGDLQDPGNCVDINECTSLLEPCRSGFSCINTVGSY	871
Qy	416	DCDPNTQASCECPBGY--ILDDGFICTDIDCEANGGFCSG---VCHNLPGTFTECIGP	468
Db	872	TCORNPVLV---CGRGYHANBESGCDVNBECETGVHRCGEOGLYCNLPUSYRCDCKP	925

Search completed: December 16, 2002, 17:25:00  
Job time : 20.5 secs



CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF136537; AAG01572.1; -  
CC EMBL: AF160978; AAF80402.1; -  
CC HSP: P35535; 1EMN.  
CC InterPro: IPR00152; ASX\_hydroxyl.  
CC InterPro: IPR00361; EGF-like.  
CC InterPro: IPR001881; EGF-Ca.  
CC InterPro: IPR001304; Lectin\_C.  
CC Pfam: PF00008; EGF\_5.  
CC Pfam: PF00059; lectin\_c; 1.  
CC SMART: SM00034; CLECT; 1.  
CC SMART: SM00181; EGF; 5.  
CC SMART: SM00179; EGF\_Ca; 5.  
CC SMART: SM00001; EGF-like; 2.  
CC SMART: SM00001; EGF-like; 2.  
CC PROSITE: PS00010; ASX\_HYDROXYL; 3.  
CC PROSITE: PS00615; C-TYPE LECTIN\_1; FALSE\_NEG.  
CC PROSITE: PS50041; C-TYPE LECTIN\_2; 1.  
CC PROSITE: PS01186; EGF\_2; 3.  
CC PROSITE: PS01187; EGF\_Ca; 3.  
CC Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;  
KW Repeat; Lectin.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 643 COMPLEMENT COMPONENT C1Q RECEPTOR.  
FT DOMAIN 24 571 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 572 592 POTENTIAL.  
FT DOMAIN 593 643 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 31 173 C-TYPE LECTIN.  
FT DOMAIN 257 298 EGF-LIKE 1.  
FT DOMAIN 299 341 EGF-LIKE 2.  
FT DOMAIN 342 381 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 382 423 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 424 462 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
FT DISULFID 261 272 BY SIMILARITY.  
FT DISULFID 268 282 BY SIMILARITY.  
FT DISULFID 284 297 BY SIMILARITY.  
FT DISULFID 303 314 BY SIMILARITY.  
FT DISULFID 308 325 BY SIMILARITY.  
FT DISULFID 327 340 BY SIMILARITY.  
FT DISULFID 346 355 BY SIMILARITY.  
FT DISULFID 351 364 BY SIMILARITY.  
FT DISULFID 366 380 BY SIMILARITY.  
FT DISULFID 386 397 BY SIMILARITY.  
FT DISULFID 393 406 BY SIMILARITY.  
FT DISULFID 408 422 BY SIMILARITY.  
FT DISULFID 428 437 BY SIMILARITY.  
FT DISULFID 443 446 BY SIMILARITY.  
FT DISULFID 448 461 BY SIMILARITY.  
FT CARBOHYD 322 322 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 498 498 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 417 417 E -> K (IN REF. 2).  
SQ SEQUENCE 643 AA; 68781 MW; 9AE4C933AD943DB6 CRC64;  
  
Query Match 18.28; Score 530.5; DB 1; Length 643;  
Best Local Similarity 29.78; Pred. No. 5.2e-28;  
Matches 166; Conservative 56; Mismatches 209; Indels 127; Gaps 28;  
  
Qy 4 VLVLGALAGLGPAPAEPOGGSCQVHDFCFALYGPATFLNASQICDGLRGLMTVR 63  
Db 8 LLLGLLGLWAGAAADSE----AVVCEGTACYTAHWKLSAAEAQHRCNENGNLATVK 63  
Qy 64 SSVA-----DVISLLN-----GDGVGRRRLWIGLQIPPGCGDPRKGLPLRGFWQVTDN 115  
Db 64 SEEARHVOEALQALLKTKAPSETKIG--KEWIGLQREKRGCTVHDL-PMKGFSSWVGGE 120

Qy 116 NTSYSRWARLDLNG--APLCGPLCVAVSAEAATVPSE--PIWEEQOCEV-----KADGFL 166  
Db 121 DTYSNWKYKASSISKRCVSLIIDLKPH--PSHLPKWHESPCGTPDAPCNIEGFL 178  
Qy 167 CEFHFPATCRPLAV-EPGAAAASVITTYGTPFAARGADQFQALPVGSSAAVAPLGLQ---- 221  
Db 179 KRFNFKGMSPLALGGPG-----QLTYTTPQATTSSLKAVPFASVANV-CGDEAESK 231  
Qy 222 ---LMTAPPGAVQGHWAREAP-----GAWDCSVENGCGEHAC-NAIPGAPRCQCPAGAAL 273  
Db 232 TNYLCKETTAGV-FHWSSGSLCVSPKFGCSFNNGQQDCDFEGGDCGRCRPGFL 290  
Qy 274 QADGRSCTASATOSNDLCEHFCVNPDPQG-----SYSCMCETGYRLAADOHRC 323  
Db 291 LDLLYTCAS-----RNPSSNPTCGGCMCHSVLSENYTCHCPRGYOLDSSQVHC 340  
Qy 324 EDVDDCILPSPQRCVNTGGFCHCVNYDLVDG---ECVEPVDPCFRANCEYOCQL 381  
Db 341 VDIDEC--EDSPCDQECINTPGGFHCEWVGQSSGSKKEACEDVDEC----- 386  
Qy 382 NQTSYLCVCAEGFAPITPHEPHRCMFCNQTACPADCDPNTQAS--CECEGYIL--DDGF 437  
Db 387 --TAAYSPCAQG-----CT-----NTDGSFYCSCKEGYIMSGEDST 420  
Qy 438 ICTDIDCEGNGFCGCVCHNLPGTFECICGPDPSALVRHIGTDCDSG-----KV 485  
Db 421 QCEDIDEC-LGNPCDTLCLINTDGSFRGCGPAGFELAPN-GVSCTRGSMFSELPARPQRE 478  
Qy 486 DGDGSGSGEPSPSPGGS 503  
Db 479 DKGDGKSTVPLTEMPGS 496  
  
RESULT 7  
FBN2\_MOUSE  
ID FBN2\_MOUSE STANDARD; PRT: 2907 AA.  
AC Q61555; Q63957;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrillin 2 precursor.  
GN FBN2 OR FBN-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95263670; PubMed=7744963;  
RA Zhang H., Hu W., Ramirez F.;  
RT "Developmental expression of fibrillin genes suggests heterogeneity  
RT of extracellular microfibrils.";  
RL J. Cell Biol. 129:1165-1176(1995).  
RN [2]  
RP SEQUENCE OF 210-317 FROM N.A.  
RX MEDLINE=94140368; PubMed=8307578;  
RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,  
RA Francke U.;  
RT "Fibrillin genes map to regions of conserved mouse/human synteny on  
RT mouse chromosomes 2 and 18.";  
RL Genomics 18:667-672(1993).  
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; AF074856; AAC63274.1; -
DR EMBL; AF081789; AAC62649.1; -
DR EMBL; AF099939; AAD47906.1; -
DR EMBL; AF099938; AAD47906.1; JOINED.
DR HSSP; P35555; IEMN.
DR MGD; MGI:106664; Ly68.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF00059; Lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_Ca; 3.
DR SMART; SM00001; EGF_like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_Ca; 3.
KW Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;
KW Repeat; Lectin.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 644 COMPLEMENT COMPONENT C1Q RECEPTOR.
FT DOMAIN 23 572 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 573 593 POTENTIAL.
FT DOMAIN 594 644 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 173 C-TYPE LECTIN.
FT DOMAIN 257 298 EGF-LIKE 1.
FT DOMAIN 299 341 EGF-LIKE 2.
FT DOMAIN 342 381 EGF-LIKE 3.
FT DOMAIN 382 423 EGF-LIKE 4.
FT DOMAIN 424 465 EGF-LIKE 5.
FT DISULFID 261 272 BY SIMILARITY.
FT DISULFID 268 282 BY SIMILARITY.
FT DISULFID 284 297 BY SIMILARITY.
FT DISULFID 303 314 BY SIMILARITY.
FT DISULFID 308 325 BY SIMILARITY.
FT DISULFID 327 340 BY SIMILARITY.
FT DISULFID 346 355 BY SIMILARITY.
FT DISULFID 351 364 BY SIMILARITY.
FT DISULFID 366 380 BY SIMILARITY.
FT DISULFID 386 397 BY SIMILARITY.
FT DISULFID 393 406 BY SIMILARITY.
FT DISULFID 408 422 BY SIMILARITY.
FT DISULFID 428 440 BY SIMILARITY.
FT DISULFID 436 449 BY SIMILARITY.
FT DISULFID 451 464 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 644 AA; 69354 MW; EB4351648BF8635A CRC64;

Query Match 18.6%; Score 543; DB 1; Length 644;
Best Local Similarity 30.0%; Pred. No. 7.9e-29;
Matches 167; Conservative 55; Mismatches 211; Indels 124; Gaps 27;

QY 5 LVILGALAGLPAPAEQPGGSCQVHDFCALYPGPATFLNASQICDGLRGLMTVRS 64
DB 9 LLILGLGQPWAG--AAASQ--AVVCEGTACTYAHGKLSAQAQRNNGNGLATVRS 64

QY 65 SVAA-----DVISLLNGDGVGRR--RLWTIGLQLPPCGDPRKRLGRLGFQWTVGNNTS 118
DB 65 EEARHVQALQTLLTKAPLEAKMGKFWIGLOREKNGCTYHDL-PMRGFSWVGGEDTA 123

QY 119 YSKWARDLNGAPLCGLPVASAAATVPSE-PIWEEQCEV-----KADGFLCEHFH 171
DB 124 YSNWYKASKSSCFFKRCVSLDLSLTPHPLPKWHESPCTGPPEAPGNSIEGFLCKNF 183

```

```

QY 172 PATCRPLAV-EPGAAAAVSIITYGTFPAARGADFOALPVGSSAAVAPLGLQ-----LM 223
DB 184 KGMCRPLALGGPG-----RVITYTTPFOATTSLEAVFPASVANVA-CGDEAKSETHYFL 236
QY 224 CT-APPGAVQGHAREAP-----GAWDCSVENGSGCEHAC-NAIPGAPRCPCPAGAAQAQDG 277
DB 237 CNEKTPGIF--HWSSGGLCVSPKFGCSFNNGGQQDCFEFGDGSFRCRPGFRLLDDL 294
QY 278 RSCTASATQSCNDLCEHCFVNPDPG-----SYSCMCETGYRLAADQHRCEVDV 327
DB 295 VTCAS-----RNPSSNPCTGGGMCHSVPLSENTCRPCSGYQLDSSVHCVDDID 344
QY 328 DCILEPSPCPQRCVNTQGGFECHCYPNYDLVG---ECVEPVDPCFRANCEYQCPLNQT 384
DB 345 EC-QDSPCAQDCVNTLGSFHCECWGYQ-PSGPKEACEDVDECAANSP----- 392
QY 385 SYLCVCAEAGFAPIPHEPHRCQMFQNTACPADCDPNTQAS--CECEGYIL--DDGFICT 440
DB 393 -----CAQGGCI-----NTDGSFYCSCKEGYIVSGEDSTQCE 423
QY 441 DIDECEN--GGFCSGVCHNLPGTFECICGPDLSALVRHIGTDCDSG-----KVD 486
DB 424 DIDECSDARGNFCDSLCTFNTDGSFRCGCPGWELAPN-GVFCSRGTVFSELPARPPQKED 482
QY 487 GGDGSGGSEPPSPPTGS 503
DB 483 NDRKESTWPTTEMPSS 499

RESULT 6
CD93_RAT STANDARD; PRT; 543 AA.
ID CD93_RAT Q9JIZ6;
AC Q9JIZ6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement component C1q receptor precursor (Complement component 1, q
subcomponent, receptor 1) (ClqR) (ClqR(p) (Clq/MBL/SPA receptor)
DE (CD93 antigen) (Cell surface antigen AA4).
GN ClQRI OR CD93 OR ClQRP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PVG; TISSUE=Natural killer cells;
RX MEDLINE=20545218; PubMed=11093152;
RA Lovik G., Vaage J.F., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;
RT "Characterization and molecular cloning of rat ClqR, a receptor on NK
cells.";
RL Eur. J. Immunol. 30:3355-3362(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Lung;
RX MEDLINE=20507883; PubMed=10934210;
RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;
RT "Molecular and cellular properties of the rat AA4 antigen, a C-type
lectin-like receptor with structural homology to thrombomodulin.";
RL J. Biol. Chem. 275:34382-34392(2000).
CC -I- FUNCTION: Receptor (or element of a larger receptor complex) for
C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
protein A (SPA). May mediate the enhancement of phagocytosis in
monocytes and macrophages upon interaction with soluble defense
collagens. May play a role in intercellular adhesion.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and
heart. Expressed at lower level in brain, thymus, liver, spleen,
intestine, kidney, adrenal gland, muscle and testis. Expressed on
endothelial cells, platelets, undifferentiated monocytes and
circulating natural killer cells.
CC -I- PTM: N- and O-glycosylated (By similarity).

```

FT	DOMAIN	302	344	EGF-LIKE 2.
FT	DOMAIN	345	384	EGF-LIKE 3.
FT	DOMAIN	385	426	EGF-LIKE 4.
FT	DOMAIN	427	468	EGF-LIKE 5.
FT	DOMAIN	594	601	POLY-LEU.
FT	DISULFID	264	275	BY SIMILARITY.
FT	DISULFID	271	285	BY SIMILARITY.
FT	DISULFID	287	300	BY SIMILARITY.
FT	DISULFID	306	317	BY SIMILARITY.
FT	DISULFID	311	328	BY SIMILARITY.
FT	DISULFID	330	343	BY SIMILARITY.
FT	DISULFID	349	358	BY SIMILARITY.
FT	DISULFID	354	367	BY SIMILARITY.
FT	DISULFID	369	383	BY SIMILARITY.
FT	DISULFID	389	400	BY SIMILARITY.
FT	DISULFID	396	409	BY SIMILARITY.
FT	DISULFID	411	423	BY SIMILARITY.
FT	DISULFID	431	443	BY SIMILARITY.
FT	DISULFID	439	452	BY SIMILARITY.
FT	DISULFID	454	467	BY SIMILARITY.
FT	CARBOHYD	325	325	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	318	318	V -> A.
FT				/FTIG-VAR_013573.
FT	CONFLICT	22	22	T -> V (IN AA SEQUENCE).
FT	CONFLICT	36	36	C -> T (IN AA SEQUENCE).
FT	CONFLICT	38	39	TA -> RI (IN AA SEQUENCE).
FT	CONFLICT	155	155	S -> N (IN REF. 1).
FT	CONFLICT	186	186	G -> A (IN AA SEQUENCE).
FT	CONFLICT	492	492	S -> A (IN AA SEQUENCE).
FT	CONFLICT	496	496	R -> Q (IN AA SEQUENCE).
FT	CONFLICT	504	504	R -> G (IN AA SEQUENCE).
FT	CONFLICT	541	541	P -> S (IN REF. 1).
SQ	SEQUENCE	652 AA:	68560 MW:	EECA0F6AC55FCAC2 CRC64:

[illegible]

Qy	487	GGDGSGE-----PPSPTPGSTTTPPA	509
Db	483	EEDKGEKGVTPRAATASPTRGECTPKA	512
RESULT 5			
CD93_MOUSE			
ID	CD93_MOUSE	STANDARD;	PRT; 644 AA.
AC	089103;		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DE	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Complement component C1q receptor precursor (Complement component 1, q		
DE	subcomponent, receptor 1) (ClqRp) (ClqR(p)) (Clq/MBL/SPA receptor)		
DE	(CD93 antigen) (Cell surface antigen AA4) (Lymphocyte antigen 68).		
GN	CLQR1 OR CD93 OR CLQRP OR LY68 OR AA4.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OC	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=129/Sv;		
RX	PubMed=1074255;		
RA	Kim T.S., Park M., Nepomuceno R.R., Palmirini G., Winokur S.,		
RA	Cotman C.A., Bengtsson U., Tenner A.J.;		
RT	"Characterization of the murine homolog of ClqR(P): identical cellular		
RT	expression pattern, chromosomal location and functional activity of		
RT	the human and murine ClqR(P).";		
RL	Mol. Immunol. 37:377-389(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Leukemia;		
RX	PubMed=10403644;		
RA	Petrenko O., Beavis A., Klaine M., Kittappa R., Godin I.,		
RA	Lemischka I.R.;		
RT	"The molecular characterization of the fetal stem cell marker AA4.";		
RL	Immunity 10:691-700(1999).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=129/Sv; TISSUE=Spleen, and Endothelial cells;		
RX	MEDLINE=99359842; PubMed=10430665;		
RA	Norsworthy P.J., Taylor P.R., Walport M.J., Botto M.;		
RT	"Cloning of the mouse homolog of the 126-kDa human Clq/MBL/SP-A		
RT	receptor, ClqRp.";		
RL	Mamm. Genome 10:789-793(1999).		
CC	-1- FUNCTION: Receptor (or element of a larger receptor complex) for		
CC	Clq, mannose-binding lectin (MBL2) and pulmonary surfactant		
CC	protein A (SPA). May mediate the enhancement of phagocytosis in		
CC	monocytes and macrophages upon interaction with soluble defense		
CC	collagens. May play a role in intercellular adhesion. Marker for		
CC	early multipotent hematopoietic precursor cells. May play a role		
CC	in cell-cell interactions during hematopoietic and vascular		
CC	development.		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	-1- TISSUE SPECIFICITY: Expressed in lung, heart and bone marrow.		
CC	Expressed at lower level in ovary, whole embryo and fetal liver.		
CC	Not detected in brain, adult liver or thymus. Highly expressed in		
CC	peritoneal cavity and bone marrow macrophages. Not detected in		
CC	epithelial cells.		
CC	-1- DEVELOPMENTAL STAGE: First detectable in day 9 embryos, in the		
CC	endocardium and vascular endothelium in the anterior part of the		
CC	embryo. Expression in endothelial cells, initially restricted to		
CC	aorta, omphalomesenteric and umbilical arteries, later extends to		
CC	subcardinal veins, intersomitic arteries and perimeural vessels.		
CC	On day 10, detectable in the entire embryo.		
CC	-1- PTM: N- and O-glycosylated (By similarity).		
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.		
CC	-1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collabora		
CC	tion between the Swiss Institute of Bioinformatics and the EMBL collabora		

```

Db 123 CHCDTGYELVDEGVDPVDFNCEYQCPVGRSEHKICAGFAVPVPGAPHKQWFC 182
Qy 409 NOTACPADCDPNTQASCEPGYIILDDGFICTDIDECENGFCGVCNHLPGTFECICGP 468
Db 183 NOTSCPADCDPHYTCRCPEGYIIDEGSTCTDINECDT-NICPGQCHNLPGTFECICGP 241
Qy 469 DSALVRHRTGDCDGGK-----DGGDSGSGPPSPPTGSLTLP-PA-VGLVHSG 516
Db 242 DSALSGQIGIDCDPTQVNEERGTPEDYG--GSGEPVSPPTGATARPSPAPAGPLHSG 297

RESULT 4
CD93_HUMAN
ID CD93_HUMAN STANDARD; PRT; 652 AA.
AC Q9NPY3; Q00274;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Complement component C1q receptor precursor (Complement component 1, q
DE subcomponent, receptor 1) (ClqR) (ClqR(p)) (Clq/MBL/SPA receptor)
DE (CD93 antigen) (CDw93).
GN C1QR1 OR CD93.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX PubMed=9047234;
RA Nepomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
RT "cDNA cloning and primary structure analysis of ClqR(p), the human
RT Clq/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";
RL Immunity 6:119-129(1997).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT ALA-318.
RX PubMed=11781389;
RA Steinberger P., Szekeres A., Willie S., Stockl J., Selento N.,
RA Prager E., Staffler G., Madic O., Stockinger H., Knapp W.;
RT "Identification of human CD93 as the phagocytic Clq receptor (ClqR)
RT by expression cloning.";
RL J. Leukoc. Biol. 71:133-140(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavridis G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Coe N.R.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A.G., Coville G.J., Deadman R., Dhali P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levaslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RA Strauberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP CHARACTERIZATION.
RX PubMed=11994479;
RA McCreel E.P., Iwakaki N., Akatsu H., Morgan B.P., Gasque P.;
RT "Human ClqR is identical with CD93 and the mni-11 antigen but does
RT not bind Clq.";
RN [6]
RL J. Immunol. 168:5222-5232(2002).
RP O-GLYCOSYLATION.
RX PubMed=10092817;
RA Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;
RT "ClqR is a heavily O-glycosylated cell surface protein involved in
RT the regulation of phagocytic activity.";
RL J. Immunol. 162:3583-3589(1999).
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for
CC Clq, mannose-binding lectin (MBL2) and pulmonary surfactant
CC protein A (SPA). May mediate the enhancement of phagocytosis in
CC monocytes and macrophages upon interaction with soluble defense
CC collagens. May play a role in intercellular adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in endothelial cells,
CC platelets, cells of myeloid origin, such as monocytes and
CC neutrophils. Not expressed in cells of lymphoid origin.
CC -!- PTM: N- and O-glycosylated.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC -!- CAUTION: Has been sometimes referred to as a collectin receptor.
CC -!- CAUTION: According to Ref.5, Clq is not a ligand for ClqR1.
CC -!- DATABASE: NAME=PRO; NOTE=PROW 3.1-6(2001); /467246456.g.htm".
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/467246456.g.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U94333; AAB53110.1; -.
DR EMBL: AL118508; CAC00597.1; -.
DR EMBL: BC028075; AAB28075.1; -.
DR HSSP: P35555; IENN.
DR MIM: 120577; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR001187; Tissue_factor.
DR Pfam: PF00008; EGF; 5.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF01108; Tissue_fac; 1.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00179; EGF_CA; 3.
DR SMART: SM00001; EGF_like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 3.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 3.
KW Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;
KW Repeat; Lectin; Polymorphism.
FT SIGNAL 1
FT CHAIN 21
FT DOMAIN 22 652
FT DOMAIN 24 580
FT TRANSMEM 591 601
FT TRANSMEM 602 652
FT DOMAIN 602 652
FT DOMAIN 632 1774
FT DOMAIN 260 301

```

```
FT CARBOHYD 494 494 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT SIMILARITY).
SQ SEQUENCE 577 AA; 61867 MW; B20E50B0FE745014 CRC64;

Query Match
Best Local Similarity 65.1%; Score 1899; DB 1; Length 577;
Matches 341; Conservative 46; Mismatches 126; Indels 8; Gaps 4;

QY 1 MLGVVLGALAGLGFAPAEPQSGSQVCHDCALYPGPATEFLNASQICDLGRHLM 60
DB 1 MLGIFFLGVLPASLGLSALAKLQPTGSCQVCHDCALYPGPATEFLNASQICDLGRHLM 60
QY 61 TVRSSVAADVISLLQSGMDLGP--WIGLQLPQGGDDPVHLGLPLRGFQWVGTDNHTSY 119
DB 61 TVRSSVAADVISLLQSGMDLGP--WIGLQLPQGGDDPVHLGLPLRGFQWVGTDNHTSY 117
QY 120 SRWRLDNLGNAPICPLGVAYSAEATVPSEPIWEQOCEKADGFLCEPHFATCRPLA 179
DB 120 SRWRLDNLGNAPICPLGVAYSAEATVPSEPIWEQOCEKADGFLCEPHFATCRPLA 179
QY 118 SRWARPNDQAPLCPGLCVTVSTATEAPGEPAAWEKPCETQGFCEFYFTASCRPLT 177
DB 118 SRWARPNDQAPLCPGLCVTVSTATEAPGEPAAWEKPCETQGFCEFYFTASCRPLT 177
QY 180 VEP-GAAAAAVSYTGPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPQVQGHWARE 238
DB 180 VEP-GAAAAAVSYTGPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPQVQGHWARE 238
QY 178 VNTDPRAHISSTYNTPFVGSADFTQLPVGSSAAVEPLGLVLCRAPPTSEGHWARE 237
DB 178 VNTDPRAHISSTYNTPFVGSADFTQLPVGSSAAVEPLGLVLCRAPPTSEGHWARE 237
QY 239 APGAWDCSVNGGCEHACNAIPGAPCQCPAGAAQADGRSCTASATQSCNDLCEHFCVP 298
DB 239 APGAWDCSVNGGCEHACNAIPGAPCQCPAGAAQADGRSCTASATQSCNDLCEHFCVP 298
QY 238 ATGAWNCSVNGGCEHACNAIPGAPCQCPAGAAQADGRSCTASATQSCNDLCEHFCVS 297
DB 238 ATGAWNCSVNGGCEHACNAIPGAPCQCPAGAAQADGRSCTASATQSCNDLCEHFCVS 297
QY 299 NPDPGSGYSCMCTGYRLAQRHCEVDVDCILEPSPCPORCVNTGGFCHCYPNYDLV 358
DB 299 NPDPGSGYSCMCTGYRLAQRHCEVDVDCILEPSPCPORCVNTGGFCHCYPNYDLV 358
QY 298 NAEVPGSYSCMCTGYRLAQRHCEVDVDCILEPSPCPORCVNTGGFCHCYPNYDLV 357
DB 298 NAEVPGSYSCMCTGYRLAQRHCEVDVDCILEPSPCPORCVNTGGFCHCYPNYDLV 357
QY 359 DGEVCEVPDPCFRANCYQCPNLTQSYLVCVCAEGFAPIPHEPRQCMFCNQACPADCD 418
DB 359 DGEVCEVPDPCFRANCYQCPNLTQSYLVCVCAEGFAPIPHEPRQCMFCNQACPADCD 418
QY 358 DGEVCEVLDPCFGNCEFCQVPTDYRCICAPGAPKDEPHKCEMFCNETSCPADCD 417
DB 358 DGEVCEVLDPCFGNCEFCQVPTDYRCICAPGAPKDEPHKCEMFCNETSCPADCD 417
QY 419 PNTQASCECPGYLIDPGLCTIDDECENGFCGVCNHLPGTFECICGSDPSALVRHIGT 478
DB 419 PNTQASCECPGYLIDPGLCTIDDECENGFCGVCNHLPGTFECICGSDPSALVRHIGT 478
QY 418 PNSTVCECPGFLDGEVSVCTDIDECGCECTSECRNFPFGVEICICGDPDTALAGISK 477
DB 418 PNSTVCECPGFLDGEVSVCTDIDECGCECTSECRNFPFGVEICICGDPDTALAGISK 477
QY 479 DCDGSKV---DGGDSGSGEPSPPTGSLTTPAVGLVHSG 516
DB 479 DCDGSKV---DGGDSGSGEPSPPTGSLTTPAVGLVHSG 516
QY 478 DCDPIPVREDTKEBEGSGEPSPPTGSLTTPAVGLVHSG 518
DB 478 DCDPIPVREDTKEBEGSGEPSPPTGSLTTPAVGLVHSG 518

RESULT 3
TRBM_BOVIN
ID TRBM_BOVIN STANDARD; PRT; 356 AA.
AC P06579;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thrombomodulin (Fetomodulin) (TW) (Fragment).
GN THBD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87067408; PubMed=3024152;
RA Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.;
RT "Characterization of a thrombomodulin cDNA reveals structural
RL similarity to the low density lipoprotein receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838(1986).
CC CC -1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
CC SCISSORS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,
CC FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF
CC THROMBIN GENERATED.
```

```
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING
CC THROMBOMODULIN.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14657; AAA30785.1; -.
CC PIR; A25918; A25918.
CC HSP; P07204; 1TMR.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_Ca.
CC Pfam; PF00008; EGF; 5.
CC SMART; SM00179; EGF_Ca; 1.
CC SMART; SM00001; EGF_Like; 3.
CC PROSITE; PS00010; ASX_HYDROXYL; 2.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 3.
CC PROSITE; PS01187; EGF_Ca; 2.
CC Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;
CC Glycoprotein; EGF-like domain.
CC NON_TER 1 1
CC DOMAIN 1 296 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 297 320 POTENTIAL.
CC DOMAIN 321 356 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 17 57 EGF-LIKE 1.
CC DOMAIN 60 98 EGF-LIKE 2.
CC DOMAIN 99 137 EGF-LIKE 3.
CC DOMAIN 139 179 EGF-LIKE 4.
CC DOMAIN 178 214 EGF-LIKE 5.
CC DOMAIN 215 254 EGF-LIKE 6.
CC DISULFID 21 32 EGF-LIKE 6.
CC DISULFID 28 41 BY SIMILARITY.
CC DISULFID 43 56 BY SIMILARITY.
CC DISULFID 64 72 BY SIMILARITY.
CC DISULFID 68 82 BY SIMILARITY.
CC DISULFID 84 97 BY SIMILARITY.
CC DISULFID 103 114 BY SIMILARITY.
CC DISULFID 110 123 BY SIMILARITY.
CC DISULFID 125 136 BY SIMILARITY.
CC DISULFID 143 152 BY SIMILARITY.
CC DISULFID 148 162 BY SIMILARITY.
CC DISULFID 164 178 BY SIMILARITY.
CC DISULFID 182 191 BY SIMILARITY.
CC DISULFID 187 199 BY SIMILARITY.
CC DISULFID 201 213 BY SIMILARITY.
CC DISULFID 219 228 BY SIMILARITY.
CC DISULFID 224 237 BY SIMILARITY.
CC DISULFID 239 253 BY SIMILARITY.
CC CARBOHYD 271 271 O-LINKED (GLYCOSAMINOGLYCAN) (BY
CC SIMILARITY).
CC
CC QUERY SEQUENCE 356 AA; 37795 MW; 29B41F07ABE1093 CRC64;
Query Match 37.5%; Score 1092.5; DB 1; Length 356;
Best Local Similarity 63.4%; Pred. No. 3.5e-65;
Matches 189; Conservative 30; Mismatches 64; Indels 15; Gaps 6;

QY 229 GAVQGHWAREAPGAWDCSVNGGCEHACNAIPGAPCQCPAGAAQADGRSCTASATQSC 288
DB 5 GTEGRWSREAPGACWVERGCGHECKGSGAGSNCLCPADAALQADGRSCGLPAEHPC 64
QY 289 NDLCEHFCVPNPDPGSGYSCMCTGYRLAQRHCEVDVDCILEPSPCPORCVNTGGF 348
DB 65 HQJCEHFC--HLHGLGNITCICAGYLAQRHCEVDVDCILEPSPCPORCVNTGGF 122
QY 349 CHCYPNYDLVDGCEVPEVDPCFRANCYQCPNLTQSYLVCVCAEGFAPIPHEPRQCMFC 408
```

Query Match		99.9%;	Score 2912;	DB 1;	Length 575;
Best Local Similarity		99.8%;	Pred. No. 1.3e-184;		
Matches 515;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	MLGVLVLGALALAGLGGFPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM	60		
DB	1	MLGVLVLGALALAGLGGFPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM	60		
QY	61	TVRSVAADVLSLLNGDGGVRRRLWTGLQLPGCCGPKRLGPLRGFWQWTDGNNNTSYS	120		
DB	61	TVRSVAADVLSLLNGDGGVRRRLWTGLQLPGCCGPKRLGPLRGFWQWTDGNNNTSYS	120		
QY	121	RWARDLNGAPLCGPGCVAVASAAEATVPSEPIWEQQCEVKADGFLCFEHPFATCRPLAV	180		
DB	121	RWARDLNGAPLCGPGCVAVASAAEATVPSEPIWEQQCEVKADGFLCFEHPFATCRPLAV	180		
QY	181	EPGAAAAASVITYGTPFAARGADFOALPVGSSAAVAPILGLQLMCTAPPGAVQGHWAREAP	240		
DB	181	EPGAAAAASVITYGTPFAARGADFOALPVGSSAAVAPILGLQLMCTAPPGAVQGHWAREAP	240		
QY	241	GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP	300		
DB	241	GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP	300		
QY	301	DQPGSYSCMCTGYRLAADOHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360		
DB	301	DQPGSYSCMCTGYRLAADOHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG	360		
QY	361	ECVEPVDPCFRANCYQCOPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCPN	420		
DB	361	ECVEPVDPCFRANCYQCOPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCPN	420		
QY	421	TOASCEPGEYILDGFICTDIDECENGFCGSGVCHNLPFTGFCICGPDPSALVRHIGTDC	480		
DB	421	TOASCEPGEYILDGFICTDIDECENGFCGSGVCHNLPFTGFCICGPDPSALVRHIGTDC	480		
QY	481	DSGKVDGDSGSGEPPSPPTPGSTILTPPAVGLVHSG	516		
DB	481	DSGKVDGDSGSGEPPSPPTPGSTILTPPAVGLVHSG	516		
RESULT 2					
TRBM_MOUSE	STANDARD;	PRT;	577	AA.	
ID	TRBM_MOUSE	STANDARD;	PRT;	577	
AC	P15306;				
DT	01-APR-1990 (Rel. 14, Created)				
DT	01-APR-1990 (Rel. 14, Last sequence update)				
DE	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Thrombomodulin precursor (Fetomodulin) (TM).				
GN	THBD.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RX	SEQUENCE FROM N.A.				
RX	MEDLINE=89128454; PubMed=2536925;				
RA	Dittman W.A., Majerus P.W.;				
RT	"Sequence of a cDNA for mouse thrombomodulin and comparison of the				
RT	predigested mouse and human amino acid sequences.;"				
RL	Nucleic Acids Res. 17:802-802(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89008498; PubMed=2844823;				
RA	Dittman W.A., Kumada T., Sadler J.E., Majerus P.W.;				
RT	"The structure and function of mouse thrombomodulin. Phorbol				
RT	myristate acetate stimulates degradation and synthesis of				
RT	thrombomodulin without affecting mRNA levels in hemangioma cells.;"				
RL	J. Biol. Chem. 263:15815-15822(1988).				
CC	-1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR				
CC	THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS				
CC	COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE				

CC	ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA	
CC	SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,	
CC	FACTOR VA AND FACTOR VIIIA, AND THEREBY REDUCES THE AMOUNT OF	
CC	THROMBIN GENERATED.	
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.	
CC	-1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING	
CC	THROMBOMODULIN (BY SIMILARITY).	
CC	-1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; X14432; CAA32597.1; -	
DR	PIR; A32001; A32001.	
DR	PIR; S08488; S08488.	
DR	HSSP; P07204; LEGT.	
DR	MGI; 98736; Thbd.	
DR	InterPro; IPR000152; Asx_hydroxyl.	
DR	InterPro; IPR000561; EGF-like.	
DR	InterPro; IPR001881; EGF-CA.	
DR	InterPro; IPR001304; Lectin_C.	
DR	InterPro; IPR001491; Thrombomoduln.	
DR	Pfam; PF00008; EGF; 4.	
DR	Pfam; PF00059; lectin_c; 1.	
DR	PRINTS; PR00907; THROMBOMODULN.	
DR	SMART; SM00034; CLECT; 1.	
DR	SMART; SM00179; EGF-CA; 1.	
DR	SMART; SM00001; EGF-like; 5.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 2.	
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.	
DR	PROSITE; PS01186; EGF_2; 3.	
DR	PROSITE; PS01187; EGF-CA; 2.	
DR	PROSITE; PS50041; C-TYPE_LECTIN_2; 1.	
KW	Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;	
KW	Glycoprotein; Signal; EGF-like domain.	
FT	SIGNAL	1 16
FT	CHAIN	17 577
FT	DOMAIN	17 517
FT	TRANSMEM	518 541
FT	DOMAIN	542 577
FT	DOMAIN	240 280
FT	DOMAIN	283 323
FT	DOMAIN	324 362
FT	DOMAIN	364 404
FT	DOMAIN	403 439
FT	DOMAIN	440 480
FT	DISULFID	244 255
FT	DISULFID	251 264
FT	DISULFID	266 279
FT	DISULFID	287 295
FT	DISULFID	291 307
FT	DISULFID	309 322
FT	DISULFID	328 339
FT	DISULFID	335 348
FT	DISULFID	350 361
FT	DISULFID	368 377
FT	DISULFID	373 387
FT	DISULFID	389 403
FT	DISULFID	407 416
FT	DISULFID	412 424
FT	DISULFID	426 438
FT	DISULFID	444 454
FT	DISULFID	449 463
FT	DISULFID	465 479
FT	CARBOHYD	113 113
FT	CARBOHYD	243 243
FT	CARBOHYD	256 256
FT	CARBOHYD	408 408



RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths C., Griffiths M.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Levaslaiho M.H., Levensha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.F., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS.  
RX MEDLINE-94029900; PubMed-8216207;  
RA Gerleitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U.,  
RA Grinnell B.W.;  
RT "Identification of the predominant glycosaminoglycan-attachment site  
in soluble recombinant human thrombomodulin: potential regulation of  
functionality by glycosyltransferase competition for serine474.";  
RL Biochem. J. 295:131-140(1993).  
RN [7]  
RP STRUCTURE BY NMR OF 389-407.  
RX MEDLINE-96007474; PubMed-7559494;  
RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;  
RT "The structure of a 19-residue fragment from the C-loop of the fourth  
epidermal growth factor-like domain of thrombomodulin.";  
RL J. Biol. Chem. 270:23366-23372(1995).  
RN [8]  
RP STRUCTURE BY NMR OF 364-407.  
RX MEDLINE-96100636; PubMed-8528067;  
RA Meininger D.P., Hunter M.J., Komives E.A.;  
RT "Synthesis, activity, and preliminary structure of the fourth  
EGF-like domain of thrombomodulin.";  
RL Protein Sci. 4:1683-1695(1995).  
RN [9]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE-95034791; PubMed-7947766;  
RA Srinivasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;  
RT "Thrombin-bound structure of an EGF subdomain from human  
thrombomodulin determined by transferred nuclear Overhauser  
effects.";  
RL Biochemistry 33:13553-13560(1994).  
RN [10]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE-96276211; PubMed-8745396;  
RA Hrabal R., Komives E.A., Ni F.;  
RT "Structural resiliency of an EGF-like subdomain bound to its target  
protein, thrombin.";  
RL Protein Sci. 5:195-203(1996).  
RN [11]  
RP STRUCTURE BY NMR OF 405-444.  
RX MEDLINE-98035729; PubMed-9367781;  
RA Sampoli Benitez B.A., Hunter M.J., Meininger D.P., Komives E.A.;  
RT "Structure of the fifth EGF-like domain of thrombomodulin: an  
EGF-like domain with a novel disulfide-bonding pattern.";  
RL J. Mol. Biol. 273:913-926(1997).  
RN [12]  
RP VARIANT TED TYR-486.  
RX MEDLINE-95111115; PubMed-7811989;  
RA Oehlin A.-K., Marlar R.A.;  
RT "The first mutation identified in the thrombomodulin gene in a  
45-year-old man presenting with thromboembolic disease.";  
RL Blood 85:330-336(1995).  
RN [13]  
RP VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.

RX MEDLINE-97341986; PubMed-9198186;  
RA Oehlin A.-K., Norlund L., Marlar R.A.;  
RT "Thrombomodulin gene variations and thromboembolic disease.";  
RL Thromb. Haemost. 78:396-400(1997).  
RN [14]  
RP VARIANT VAL-473.  
RX MEDLINE-9720518; PubMed-9157575;  
RA Norlund L., Holm J., Zoller B., Oehlin A.-K.;  
RT "A common thrombomodulin amino acid dimorphism is associated with  
myocardial infarction.";  
RL Thromb. Haemost. 77:248-251(1997).  
RN [15]  
RP VARIANT THR-43.  
RX MEDLINE-99057299; PubMed-9843165;  
RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,  
RA Stubbs P.J., Manger Cats V., Ireland H.;  
RT "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,  
and the risk of myocardial infarction in men.";  
RL Thromb. Haemost. 80:743-748(1998).  
RN [16]  
RP VARIANT VAL-473.  
RX MEDLINE-21143723; PubMed-11245641;  
RA Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,  
RA Juneja H.;  
RT "Thrombomodulin Ala455Val polymorphism and risk of coronary heart  
disease.";  
RL Circulation 103:1386-1389(2001).  
CC [-] FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR  
THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS  
COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE  
ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA  
SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,  
FACTOR VA AND FACTOR VIIIA, AND THEREBY REDUCES THE AMOUNT OF  
THROMBIN GENERATED.  
CC [-] SUBCELLULAR LOCATION: Type I membrane protein.  
CC [-] TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING  
THROMBOMODULIN.  
CC [-] POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED  
RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).  
CC [-] DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,  
ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE  
DEFECTS OF THE HAEMOPOIETIC SYSTEM WHICH CREATES A TENDENCY TO THE  
OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE  
PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.  
CC [-] SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
CC [-] DATABASE: NAME=PROW; NOTE=cd guide cd141 entry;  
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd141.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X05495; CAA29045.1; -  
DR EMBL; M16552; AAB59508.1; -  
DR EMBL; J02973; AAA61175.1; -  
DR EMBL; D00210; BAA00149.1; -  
DR EMBL; AL049651; CAB51954.1; -  
DR PIR; A27073; A27073.  
DR PIR; A28307; A28307.  
DR PIR; A29680; A29680.  
DR PDB; 1EGT; 15-NOV-95.  
DR PDB; 1FGD; 20-JUN-96.  
DR PDB; 1FGE; 20-JUN-96.  
DR PDB; 1TMR; 08-JUN-95.  
DR PDB; 1ZAO; 29-JAN-96.  
DR PDB; 1ADX; 24-DEC-97.  
DR PDB; 2ADX; 24-DEC-97.  
DR GlycoSuiteDB; P07204; -  
DR Genex; HGNC:11784; THBD.

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:20:02 ; Search time 10.5 Seconds  
(without alignments)  
2038.265 Million cell updates/sec

Title: US-09-509-994-1  
Perfect score: 2916  
Sequence: 1 MLGVLVGLALAGLGPAP.....PSPTPGSTLTTPPAVGLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2912	99.9	575	1 TRBM_HUMAN	P07204 homo sapien
2	1899	65.1	577	1 TRBM_MOUSE	P15306 mus musculus
3	1092.5	37.5	356	1 TRBM_BOVIN	P08579 bos taurus
4	583	20.0	652	1 CD93_HUMAN	Q9npv3 homo sapien
5	543	18.6	644	1 CD93_MOUSE	O89103 mus musculus
6	530.5	18.2	643	1 CD93_RAT	Q9et61 rattus norv
7	373	12.8	2907	1 FBN2_MOUSE	G61555 mus musculus
8	367.5	12.6	1184	1 FBL2_HUMAN	P98095 homo sapien
9	367	12.6	2871	1 FBN1_BOVIN	P98133 bos taurus
10	366	12.6	2871	1 FBN1_PIG	Q9tv36 sus scrofa
11	365	12.5	2911	1 FBN2_HUMAN	P35556 homo sapien
12	363	12.4	2871	1 FBN1_HUMAN	P35555 homo sapien
13	361	12.4	2871	1 FBN1_MOUSE	G61554 mus musculus
14	358	12.3	956	1 MYN2_HUMAN	O00339 homo sapien
15	357	12.2	1221	1 FBL2_MOUSE	P37889 mus musculus
16	329.5	11.3	1964	1 NTC4_MOUSE	P31695 mus musculus
17	328.5	11.3	956	1 MTN2_MOUSE	O08746 mus musculus
18	327.5	11.2	1712	1 LTB1_RAT	Q00918 rattus norv
19	324.5	11.1	712	1 FBL1_CAEEL	O77469 caenorhabdi
20	320.5	11.0	2321	1 NTC3_HUMAN	Q9um47 homo sapien
21	320	11.0	1394	1 LTB5_HUMAN	P22064 homo sapien
22	320	11.0	1595	1 LTB1_HUMAN	Q14766 homo sapien
23	315.5	10.8	443	1 FBL4_HUMAN	O95967 homo sapien
24	313	10.7	703	1 FBL1_HUMAN	P23142 homo sapien
25	308.5	10.6	443	1 FBL4_CRIGR	O55058 cricetus
26	308.5	10.6	2003	1 NTC4_HUMAN	Q99466 homo sapien
27	308	10.6	448	1 FBL5_HUMAN	Q9ubx5 homo sapien
28	308	10.6	705	1 FBL1_MOUSE	Q08879 mus musculus
29	307	10.5	684	1 FBL1_CHICK	O73775 gallus gall
30	306.5	10.5	443	1 FBL4_MOUSE	Q9wv19 mus musculus
31	302	10.4	448	1 FBL5_MOUSE	Q9wv19 mus musculus
32	296	10.2	448	1 FBL5_RAT	Q9wv18 rattus norv
33	296	10.2	2531	1 NTC1_MOUSE	Q01705 mus musculus

RESULT 1

ID	TRBM_HUMAN	STANDARD;	PRT;	575 AA.
AC	P07204;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen).			
GN	THBD OR THRM.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88004395; PubMed=2820710;			
RA	SUZUKI K., Kusumoto H., Dayashiki Y., Nishioka J., Maruyama I.,			
RA	Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.;			
RT	"Structure and expression of human thrombomodulin, a thrombin			
RT	receptor on endothelium acting as a cofactor for protein C			
RT	activation."			
RL	EMBO J. 6:1891-1897(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88024950; PubMed=2822087;			
RA	Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.;			
RT	"Human thrombomodulin: complete cDNA sequence and chromosome			
RT	localization of the gene."			
RL	Biochemistry 26:4350-4357(1987).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87317665; PubMed=2819876;			
RA	Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.;			
RT	"Human thrombomodulin gene is intron depleted: nucleic acid sequences			
RT	of the cDNA and gene predict protein structure and suggest sites of			
RT	regulatory control."			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88227901; PubMed=2836377;			
RA	Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,			
RA	Dayashiki Y., Maruyama I., Suzuki K.;			
RT	"Gene structure of human thrombomodulin, a cofactor for thrombin-			
RT	catalyzed activation of protein C."			
RL	J. Biochem. 103:281-285(1988).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			

34	293.5	10.1	2319	1	NTC3_RAT	Q9r172 rattus norv
35	291	10.0	2470	1	NTC2_MOUSE	O35516 mus musculus
36	288.5	9.9	493	1	FBL3_HUMAN	Q12805 homo sapien
37	286	9.8	2471	1	NTC2_RAT	Q9gw30 rattus norv
38	284	9.7	2471	1	NTC2_HUMAN	Q04721 homo sapien
39	283.5	9.7	2437	1	NTC1_BRARE	P46530 brachydanio
40	280.5	9.6	2703	1	NTC3_MOUSE	P07207 drosophila
41	280	9.6	2318	1	NTC3_MOUSE	Q61982 mus musculus
42	277	9.5	2531	1	NTC1_RAT	Q07008 rattus norv
43	276.5	9.5	493	1	FBL3_RAT	O35568 rattus norv
44	271	9.3	2556	1	NTC1_HUMAN	P46531 homo sapien
45	269.5	9.2	1217	1	EGF_MOUSE	P01132 mus musculus

ALIGNMENTS

Search completed: December 16, 2002, 17:26:48  
Job time : 25.5 secs

Db	580	LGTEAELPNLPGDDQDECIMLPCELCQHLICINTVGSYRCACFPFGELQGDGTCRPRDGR	639
QY	230	AVGQHWARE--APGAWDCSV-----ENGCEHACNAIPGAPRCQCPAGAA	272
Db	640	APQLDARESAPRSESAQVSPNTIPDPVQPNCTCKONGPCRCRVVGDTAMCSCFPGYA	699
QY	273	LOADGRSC-----TASATQSCNDLCBHFCHVPNPDPQSYSC-----MCEYTRFLAADQHR	322
Db	700	IMADGVSCDEQDECLMTHDCS--WKQFCV---NTLGSFVVCVNHVTLCAEYILNA--HRK	753
QY	323	CEYDDCILEPSPSC--PQRCVNTQGGFECH-----CYPNTYDLVDGCEVPEVDPCFRA--NC	374

```

Db      754  CVDINECVTLHTCTTAEHCNVTNPGSFQCYKALTCEPGVYLTDECTD-VDECVTGTHNC 812
      QY      375  E-YQCOPLNQTSVLCV-----CAEGFAPIPH-----EPRHQMECNQACPA 415
      Db      813  QAQFSCQN-TKGIFYCOARQRCMDGFLQDPGNCVINECTSLLEPCRGFSFCINTVGSY 871
      QY      416  DCDPNTQASCEPEGY-ILDDGFICITDIDCEGNGFCSG---VCHNLPPTFEICGP 468
      Db      872  TCQRNPLV---CGRGYHANESEGVNECEGTGVRHCEGGQLCYNLPGSYRCDKRP 925

RESULT 14
T43210
fibulin-LD precursor - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C:Accession: T43210
R:Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.
submitted to the EMBL Data Library, June 1998
A:Description: Identification of chicken and C. elegans fibulin-1 homologs
A:Reference number: Z29337
A:Accession: T43210
A>Status: preliminary; translated from GB/EMBL/DDHJ

```

A;Residues: 1-569 <BAK>  
A;Cross-references: EMBL:AF070477; PIDN:AAC24035.1  
C;Genetics:  
A;Note: Intron positions not resolved (incomplete sequence)  
C;Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 11.4%; Score 331; DB 2; Length 589;  
Best Local Similarity 28.1%; Pred.No. 5.9e-13;  
Matches 126; Conservative 49; Mismatches 166; Indels 108; Gaps

QY 103 GPLRGF-QWTTGDNNTSYSRWARLDNLGAPLCGPLCAVVASAAATVPSEPI-----WEEQ 156  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 19 GCLRSFNKCNGDIEITH---ASEITGRPLNDPHVLHLGDRCASSHCHELCHDRGEKV 75  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
QY 157 QCEVKA-----DGFLCEHFHP--ATCRPLAVE-----PGAAAAAVSYITYGTPEA--- 198  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 76 ECSCRSGFDLAPGMCAVDHIDCATILMDDCLESQRCLNTPGSFKICRILTSCGTGYAMDS 135  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
QY 199 --ARGADFOALPYGVSSAAVPLGLQLMCTAPPQAVQ-----GHWAREAPAWDCS---- 246  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 136 ETERCIRDVEDCNLGSH-----DCGPLYCRRNTQCSYRCDAKKCGDGLQNPMTGCETISITC 191  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
QY 247 -----VENGGCE-----HACNA-----IFGARPCQ-----CPAGAAALQADGRSCTA 282  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 192 PNGYIPKNGMCNIDECVTGHNGCAGECVNTGPSFRCCQOKNLCAHGEVN----- 243

[illegible]

R:Perreira, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangillan, T.; Bonad  
Hum. Mol. Genet. 2, 961-968, 1993  
A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene B  
A:Reference number: I54355; MUID:93372860; PMID:8364578  
A:Accession: I54355  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 132-3002 <PER>  
A:Cross-references: GB:I13923; NID:9306745; PIDN:AAB02036.1; PID:g306746  
R:Maslen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.  
Nature 352, 334-337, 1991  
A:Title: Partial sequence of a candidate gene for the Marfan syndrome.  
A:Reference number: S17064; MUID:91304568; PMID:1852207  
A:Accession: S17064  
A:Molecule type: mRNA  
A:Residues: 1030-3002 <MAS>  
A:Cross-references: EMBL:X63556  
R:Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.  
Science 259, 680-683, 1993  
A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.  
A:Reference number: I59574; MUID:93157831; PMID:8430317  
A:Accession: I59574  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 2217-2288, 'I', 2290-2325 <RES>  
A:Cross-references: GB:S54426; NID:9264860; PIDN:AAB25244.1; PID:g264861  
R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.  
Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differe  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17062  
A:Molecule type: mRNA  
A:Residues: 'VLTVTFVIFISYNKML', 944-1444 <LEE1>  
A:Cross-references: EMBL:X62008; NID:g31398; PIDN:CAB56534.1; PID:g5924015  
A:Accession: S62111  
A:Molecule type: protein  
A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>  
R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.  
J. Biol. Chem. 264, 21381-21385, 1989  
A:Title: Connective tissue microfibrils. Isolation and characterization of three large F  
A:Reference number: A34198; MUID:90078246; PMID:2512293  
A:Accession: A34198  
A:Molecule type: protein  
A:Residues: 565-575, 1890-1892, 'I', 1894-1900 <MAD>  
C:Comment: Fibrillin is a major component of elastin-associated microfibrils.  
C:Genetics:  
A:Gene: GDB:FBN1  
A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700  
A:Map position: 15q21.1-15q21.1  
A:Introns: 2236/1; 2258/1; 2297/1  
C:Superfamily: fibrillin 1; EGF homology  
C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; M  
F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predicted  
F:1332-1367/Domain: fibrillin (5'-region exon C splice form) #status predicted <MATC>  
F:1457-1492/Domain: EGF homology <EGF>  
F:2262-2295/Domain: EGF homology <EGF1>  
Query Match 12.4%; Score 363; DB 2; Length 3002;  
Best Local Similarity 30.8%; Pred. No. 2.8e-14;  
Matches 111; Conservative 22; Mismatches 105; Indels 122; Gaps 20;  
Qy 244 DCSVENGGCEHACNAIFGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP- 300  
Db 1331 ECSIMNGGCTCTNSGSEYCSQPGFALMPDQSRCT-----DIDE--CEDNPNIC 1380  
Qy 301 -----DQPGSYSCMCTGTVLAADQHRCEVDVDCILEPSPC-PQRCVNTQGGFECHECYP 353  
Db 1381 DGGQCTNIPGEYRCLCYVDGFWASDMKTCVDVNECDLNPICLSGTCTKGTGFIHCDM 1440  
Qy 354 NYDLVDGE--CPEVPDPC--FRANCEYQCQPLNQT-SYLCVCAEGFA-----PIP 398  
Db 1441 GYSKKKGTGCTD-INECEIGHNCGKHAVCTNTAGSFKCSGSPGWIGDGKICTDLDECS 1499

Qy 399 HEPHRCQMFNCQTACPADCDNNTQAS--CECEGYLDGDFICTDIDEC-EN----- 447  
Db 1500 NTHMCSQH-----ADC-KNTMGSYRCLCKEGY-TGDGFTCTDLDECSSENLCGNGQ 1550  
Qy 448 -----GGF-----CS-----GVCHNLPGTFECICGPDSSALV 473  
Db 1551 CLNAPGGYRCCECDMGFVPSADGKACEDIDECSLPNICVFGTCHNLPGFRCCECEGYELD 1610  
Qy 474 RHIG-----TDCDSGKVDGSGSGEPSPPTPGS-----TLTPPAVGLV 513  
Db 1611 RSGGCTDVNECLDPTTCTISGNCVN-----TPGSYICDPPDFELNPTRVGCV 1658  
RESULT 11  
A55624  
fibrillin-1 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 02-Aug-2002  
C:Accession: A55624  
R:Yin, W.; Smiley, E.; Germiller, J.; Sanguinetti, C.; Lawton, T.; Perreira, L.; Ramire  
J. Biol. Chem. 270, 1798-1806, 1995  
A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin  
A:Reference number: A55624; MUID:95130561; PMID:7829516  
A:Accession: A55624  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2871 <YIN>  
A:Cross-references: GB:L29454; NID:g575509; PIDN:AAA56840.1; PID:g575510  
C:Genetics:  
A:Gene: Fbn-1  
A:Superfamily: fibrillin 1; EGF homology  
F:1201-1236/Domain: EGF homology <EGF>  
Query Match 12.4%; Score 361; DB 2; Length 2871;  
Best Local Similarity 29.4%; Pred. No. 3.5e-14;  
Matches 117; Conservative 26; Mismatches 127; Indels 138; Gaps 21;  
Qy 212 SAAVAPLGL-----QLMCTAPGAVOGHWAREAGWDSCVSENGGCEHACNAIFGAPRC 265  
Db 1162 SANLCPHGRVNLIGKYQACNPGYHPTDRLFCVDIDECSIMNGGCTCTNSGSEYEC 1221  
Qy 266 QCPAGAAQADGRSCTASATQSCNDLCEHFCVNP------DQPGSYSCMCTGYRL 316  
Db 1222 SQPGFALMPDQSRCT-----DIDQ--CEDNPNICDGGQCTNIPGEYRCLCYVDGFWA 1271  
Qy 317 AADQHRCEVDVDCILEPSPC-PQRCVNTQGGFECHECYPNYDLVDGE--CPEVPDPC--FR 371  
Db 1272 SEMDKTCVDVNECDLNPICLSGTCTKGTGFIHCDMGTGKGTGCTD-INECEIGA 1330  
Qy 372 ANCEYQCQPLNQT-SYLCVCAEGFA-----PIPEHRCQMFNCQTACPADCDN 420  
Db 1331 HNGRHAVCTNTAGSFKCSGSPGWIGDGKICTDLDECSNGTHMCSQH-----ADC-KN 1382  
Qy 421 TOAS--CECEGYLDGDFICTDIDEC-EN-----GGF----- 450  
Db 1383 TMGSYRCLCKGY-TGDGFTCTDLDECSSENLCGNGQCLNAPGGYRCCECDMGFVPSADG 1441  
Qy 451 -----CS-----GVCHNLPGTFECICGPDSSALVRHIG-----TDCDSK 484  
Db 1442 KACEDIDECSLPNICVFGTCHNLPGFRCCECEGYELDRSGGNCITDVNECLDPTTCTISGN 1501  
Qy 485 VDGDSGSGEPSPPTPGS-----TLTPPAVGLV 513  
Db 1502 CVN-----TPGSYTCDCSPDFELNPTRVGCV 1527  
RESULT 12  
T46488  
hypothetical protein DKFp434J065.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C:Accession: T46488

```

Db      683 CKDNGCKQVC---STVGGSATCSFPGYAIWAGVSCDINECVTOLHTCSRGEHCVNT 739
QY      344 QGGFECH---CYPNVDLVGCEVPDPFCFRANCEYQCQPLNLTSLCV----- 389
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      740 LGSFHCYKALTCFPGYALKDGC-EDVDEC--AMGHTHCQP---GFLCQNTKGSFYCQA 792
QY      390 ---CAGFPAPIH-----EPHRCQMFNCWTACPADCPDNTQASCCEPQY-I 432
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      793 RQRCMDGFLQDPGNCVDINECTSLSEPCRFGFCINTVGSYTCQRNPLI---CARGYHA 849
QY      433 LDPGFICTIDCENGGFCSG---VCHNLPGTFCEIC 466
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      850 SDDGAKVDVNECEGTGVRHCGEQVCHNLPGSYRDC 886
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
A55567
fibrillin 1 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Aug-2002
C:Accession: A55567
R:Tilstra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.
Genomics 23, 480-485, 1994
A:title: Sequence of the coding region of the bovine fibrillin cDNA and local-
A:Reference number: A55567; MUID:95137597; PMID:7835900
A:Accession: A55567
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <TIL>
A:Cross-references: GB:L28748; NID:g508427; PIDN:AAA74122.1; PID:g508428
C:Superfamily: fibrillin 1; EGF homology
F:1201-1236/Domain: EGF homology <EGF>

Query Match 12.6%; Score 367; DB 2; Length 2871;
Best Local Similarity 31.5%; Pred. No. 1.5e-14;
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps

QY      244 DCSVNGGCGEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCHEFCVNP- 300
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1200 ECSIMGGCTETCTNSEGSYECQPGFALMPDQRSC-----DIDE--CEDNPNIC 1249
QY      301 -----DQPGSYSCMCETGYRLAADQHRCEDVDDCIILEPSPC-PQRCVNTQGGFECHCYP 353
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1250 DGGGCTNIPGEYRCLCYDGFMASEDWKTCVDVNECDLNPNICLSGTCENTKGSFICHDM 1309
QY      354 NYDLVGE--CYEVPDPC--FRANCEYQCQPLNQT-SYLCVCAAGFA-----PIP 398
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1310 GYSGKAGKGTCTD-INECEIGAHNCDRHAVCNTINTAGSKFSCSPGWDGDKTCDLDECS 1368
QY      399 HEPHRCQMFNCQTACPADCPDNTQAS--CECPGYLLDDGFICTDIDEC-EN----- 447
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEGY-TGDGFTCTDLDECSNLMLCNGQ 1419
QY      448 ---GGF-----CS-----GVCHNLPGTFCEICGPDVALV 473
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1420 CLNAPGGYRCBDMGFVPSADGKACEDIDCSLPNICVFGTCHNLPLGLFRCEELGYELD 1479
QY      474 RHIG-----TDCDSGK-VDGGSGSGGEPPTPTGTTTPPAVLV 513
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1480 RSGGNCCTDVNECLDPTTCLISGGCNVTPGYSYCDCCPPD----FELNPRVGCV 1527

```

[illegible]

C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: T13954  
R:Nakajima, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A:Reference number: Z14126; MUID:98360089; PMID:9693030  
A:Accession: T13954  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1574 <NAK>  
A:Cross-references: EMBL:AB011532; NID:g3449293; PIDN:BAA32462.1; PID:g3449294  
A:Experimental source: strain Sprague-Dawley; brain  
C:Genetics:  
A:Gene: MEGF6

Query Match 14.2%; Score 414; DB 2; Length 1574;  
Best Local Similarity 34.1%; Pred. No. 1.3e-17;  
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 17;

QY 224 CTAPPG-AVGHWAREAPGAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTA 282  
Db 147 CRPPGVQLQGD-GKTCQDVDECAHNGGQHRCVNTPGSYLCECKPGRFLHTDGRCL- 204  
QY 283 SATQSC---NDLCEHFCVP-----NPDQGSYSYSCM----- 309  
Db 205 -AISCTLNGGCGHQCQVLTQTHRCQCRPQYQLQEDGRRVCRRSPCAEGNGGCMHICQ 263  
QY 310 -----CETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGECHECHYPNYDL-VD 359  
Db 264 ELRLAHRCGHPGYQLAADRKTCDVDDECAHGLAQAACGCLNTQGSFKVCYCHAGYELGAD 323  
QY 360 G-EG---VEPVDPCFRAN---CEVQCPQLNTSYLVCACAEGFAPIPHEPHRCQMF---- 408  
Db 324 GRQCYRTEIIVNCEAGNGGCSHGCSH-TSTGFLCTCPRGY-----ELDEDQKTCIDID 377  
QY 409 ---NOTACPADCDNTQA---SCEPEGYILD-DGFICTIDIDCEBNG-GFCSGVGNLPGT 461  
Db 378 DCANSPCCQAC-ANTPGYECSCFAGYRLNTDGGCEDVDGASGGHGCHEHCSNLAGS 436  
QY 462 FECIC 466  
Db 437 FOCFC 441

RESULT 6  
A57278  
fibullin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 02-Aug-2002  
C:Accession: A57278  
R:Zhang, H.; Hu, W.; Ramirez, F.  
J. Cell Biol. 129, 1165-1176, 1995  
A:Title: Developmental expression of fibullin genes suggests heterogeneity of extracellular matrix  
A:Reference number: A57278; MUID:95263670; PMID:7744963  
A:Accession: A57278  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2907 <ZHA>  
A:Cross-references: GB:L39790; NID:g762830; PIDN:AAA74908.1; PID:g762831  
C:Superfamily: fibullin 1; EGF homology  
F:1239-1274/Domain: EGF homology <EGF>  
F:2488-2523/Domain: EGF homology <EGF>

Query Match 12.8%; Score 373; DB 2; Length 2907;  
Best Local Similarity 32.9%; Pred. No. 6.8e-15;  
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

QY 244 DCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDP-- 301  
Db 1238 ECMINMGCDTQCTNSEGSYECSCSEGYALMPDGRSCA-----DIDE--CENNPDIC 1287  
QY 302 -----QPGSYSCMCTGYRLAADOHRCEVDVDCILEPSPCP-QRCVNTQGGFECCHCYP 353

Db 1288 DGGQCTNIPGEYRCLCYDGFMAHMDKTCIDVNECDLNPNICMFCGECENTKGSFICHQCL 1347  
QY 354 NYDLVDGE--CPEVDPC--FRANCEYQCPQLN-QTSYLCVCAEGFA-----PIP 398  
Db 1348 GYSVKKCTTGCTD-VDECEIGAHCNDMHASCLNVPVGFKSCREGVWGVNGIKICIDLDECA 1406  
QY 399 HEPHRCQMFQNTACPADCDNTQA--CECEGYILLDDGFICTIDID-----CENG- 448  
Db 1407 NGTHQCSI-----NAQC-VNTPGSYRCACSEGF-TGDGFTCSVDVDECAENTNLCENGQ 1457  
QY 449 -----GF-----CS-----GVCHNLPGTFECICGPDLSALV 473  
Db 1458 CLNVPGAYRCCEMGFTPASDSRSCQIDECFQNICVFGTCNNLPGMFHCICDDGYELD 1517  
QY 474 RHIG--TDGD 481  
Db 1518 RTGGNCTDID 1527

RESULT 7  
A55184  
fibullin-2 precursor - human  
N:Alternate names: protein DKFzp586A1519.1  
C:Species: Homo sapiens (man)  
C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 02-Aug-2002  
C:Accession: A55184; T08744  
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.  
Genomics 22, 425-430, 1994  
A:Title: Fibullin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene  
A:Reference number: A55184; MUID:95104855; PMID:7806230  
A:Accession: A55184  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1184 <ZHA>  
A:Cross-references: GB:X82494; NID:g575232; PIDN:CAA57876.1; PID:g575233  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z16471  
A:Accession: T08744  
A:Molecule type: mRNA  
A:Residues: 656-719; QDECLMGADCSRROFCVNTLGSFYCVNHTVLCADGYILNAHRKCDV', 720-853, 'T', 855  
A:Cross-references: EMBL:AL050095  
A:Experimental source: adult uterus; clone DKFzp586A1519  
C:Genetics:  
A:Gene: GDB:FBLN2  
A:Cross-references: GDB:293037; OMIM:135821  
A:Map position: 3p25-3p24  
A:Note: DKFzp586A1519.1  
C:Superfamily: fibullin-2; EGF homology  
C:Keywords: alternative splicing; extracellular matrix  
F:1-27/Domain: signal sequence status predicted <SIG>  
F:28-1184/Product: fibullin-2 protein #status predicted <MAT>  
F:905-941/Domain: EGF homology <EGF>

Query Match 12.6%; Score 367.5; DB 2; Length 1184;  
Best Local Similarity 30.0%; Pred. No. 6.7e-15;  
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;

QY 154 EEQOCEVKAD-GFLCEFHFPATC-----RPLAV-----EPGAAAASVTYGTFFAARG 201  
Db 531 EGQSCSNPLGYPCN-HVMLSCCEGEPLIVPEVRRPPPEAAAPRRVS-----EAEM 582  
QY 202 AFQALPVGSSAAV---APLGQLMCTAPGAVQGHAREAPGAWDCSVENGCEHACNA 258  
Db 583 AGREALSLGTEALPNSLPGLDQDDECLLPGL-----CQHLICIN 622  
QY 259 IFCAPRCQCPAGALQADGRSC-----TASAT-----QS 287  
Db 623 TVGSYHCACFPFGLSDGGDGTCTCPGHPPOPEAPQEPALKSEFSQVANSNIPPLPQNPNT 682  
QY 288 CND--LCEHFCVNPDPQGSYSCMCTGYRLAADOHRCEVDVDCILEPSPCP--QRCVNT 343



[illegible]

RESULT 3

A25918  
thrombomodulin - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 16-Jul-1999  
C:Accession: A25918  
R:Jackman, R.W.; Beeler, D.L.; VanDeWater, L.; Rosenberg, R.D.  
Proc. Natl. Acad. Sci. U.S.A. 83, 8834-8838, 1986  
A:Title: Characterization of a thrombomodulin cDNA reveals structural similarity to the  
A:Reference number: A25918; MUID:87067408; PMID:3024152  
A:Accession: A25918  
A:Molecule type: mRNA  
A:Residues: 1-356 <JAC>  
A:Cross-references: GB:M14557; NID:gi63762; PIDN:AAA30785.1; PID:gi63763  
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C:Keywords: transmembrane protein  
F:21-56/Domain: EGF homology <EG1>  
F:64-97/Domain: EGF homology <EG2>  
F:103-136/Domain: EGF homology <EG3>  
F:143-178/Domain: EGF homology <EG4>  
F:182-213/Domain: EGF homology <EG5>  
F:219-253/Domain: EGF homology <EG6>

	Query Match	37.5%	Score 1092.5	DB 2	Length 356
	Best Local Similarity	63.4%	Pred. No. 4.1e-59		
	Matches 189	Conservative 30	Mismatches 64	Indels 15	Gaps
QY	229	GA VQGHWAREAPAGWDCSVENGCEHACNAIFCAPQCQAGAAQAQDGRSTASATQSC	288		
	:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :				
Db	5	GETEGWRSREAPGAWACGVGEGCQHECKGSAGASNCLCPADAAALQADRCGLPAEHPC	64		
QY	289	NDLCEHFVCPNPDQCSYSCWCBTGYRLAADQHRCEVDVDCILFSPQPCQRCVNTQGGFE	348		
	:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :				
Db	65	HQLCEHFC--HLHGLGNVTCICEAGYQLAADQHRCEVDVDCIQLFSPQPCQRCVNTQGGFQ	122		

```

QY   349  CHCYPNYDLVDSCEVPDPFRANCEYQCQLNNTSYLCVCAEGFAPITPIEHPRCMQFC  408
      ||| ||||| ||||| ||||| ||||| : : |:||||| : |:|||||
Db    123  CHCDTGYELVDGECVDPDPCFDNNNCEYQCQPVRSEHKICIAEGFAPVPGAPHKCMQFC  182

QY   409  NQTACPADCDPNTQAACEPEGYILDDGFICHTIDBCENGGCSCGVCHNLPTCFECIGCP  468
      ||| ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db    183  NQTSACPADCDDPHYPTTCRPEGYIIDEGSTCTDINECDT-NICPGQCHNLPTGYECIGCP  241

QY   469  DSALVRHIGTDCDSKV-----DGGDSGSGEPPSPPTPGSTLTP-PA-VGLVHSG  516
      ||| ||| ||| | | | | | | | | | | | | | | | | | | | |
Db    242  DLSALSGQIGIDCDPTQVNEERTPEDYG--GSGEPPSVPTGTATARPSPAPAGPLHS  297

RESULT 4
T27283
hypothetical protein Y64G10A.f - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27283
R:Ainscough, R.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z20336
A;Accession: T27283
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1620 <WILL>
A;Cross-references: EMBL: AL110498; NID: el542303; PIDN: CAB54471.1; CESP: Y64G10A.f
A;Experimental source: clone Y64G10A
```

```

RESULT 4
T27283
hypothetical protein Y64G10A.f - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27283
R:Ainscough, R.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20336
A:Accession: T27283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1620 <WIL>
A:Cross-references: EMBL:AL110498; NID:el542303; PIDN:CAB54471.1; CESP:Y64G10A.f
A:Experimental source: clone Y64G10A
C:Genetics:
A:Gene: CESP:Y64G10A.f
A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 604/1; 605/1; 606/1; 607/1; 608/1; 609/1; 610/1; 611/1; 612/1; 613/1; 614/1; 615/1; 616/1; 617/1; 618/1; 619/1; 620/1; 621/1; 622/1; 623/1; 624/1; 625/1; 626/1; 627/1; 628/1; 629/1; 630/1; 631/1; 632/1; 633/1; 634/1; 635/1; 636/1; 637/1; 638/1; 639/1; 640/1; 641/1; 642/1; 643/1; 644/1; 645/1; 646/1; 647/1; 648/1; 649/1; 650/1; 651/1; 652/1; 653/1; 654/1; 655/1; 656/1; 657/1; 658/1; 659/1; 660/1; 661/1; 662/1; 663/1; 664/1; 665/1; 666/1; 667/1; 668/1; 669/1; 670/1; 671/1; 672/1; 673/1; 674/1; 675/1; 676/1; 677/1; 678/1; 679/1; 680/1; 681/1; 682/1; 683/1; 684/1; 685/1; 686/1; 687/1; 688/1; 689/1; 690/1; 691/1; 692/1; 693/1; 694/1; 695/1; 696/1; 697/1; 698/1; 699/1; 700/1; 701/1; 702/1; 703/1; 704/1; 705/1; 706/1; 707/1; 708/1; 709/1; 710/1; 711/1; 712/1; 713/1; 714/1; 715/1; 716/1; 717/1; 718/1; 719/1; 720/1; 721/1; 722/1; 723/1; 724/1; 725/1; 726/1; 727/1; 728/1; 729/1; 730/1; 731/1; 732/1; 733/1; 734/1; 735/1; 736/1; 737/1; 738/1; 739/1; 740/1; 741/1; 742/1; 743/1; 744/1; 745/1; 746/1; 747/1; 748/1; 749/1; 750/1; 751/1; 752/1; 753/1; 754/1; 755/1; 756/1; 757/1; 758/1; 759/1; 760/1; 761/1; 762/1; 763/1; 764/1; 765/1; 766/1; 767/1; 768/1; 769/1; 770/1; 771/1; 772/1; 773/1; 774/1; 775/1; 776/1; 777/1; 778/1; 779/1; 780/1; 781/1; 782/1; 783/1; 784/1; 785/1; 786/1; 787/1; 788/1; 789/1; 790/1; 791/1; 792/1; 793/1; 794/1; 795/1; 796/1; 797/1; 798/1; 799/1; 800/1; 801/1; 802/1; 803/1; 804/1; 805/1; 806/1; 807/1; 808/1; 809/1; 810/1; 811/1; 812/1; 813/1; 814/1; 815/1; 816/1; 817/1; 818/1; 819/1; 820/1; 821/1; 822/1; 823/1; 824/1; 825/1; 826/1; 827/1; 828/1; 829/1; 830/1; 831/1; 832/1; 833/1; 834/1; 835/1; 836/1; 837/1; 838/1; 839/1; 840/1; 841/1; 842/1; 843/1; 844/1; 845/1; 846/1; 847/1; 848/1; 849/1; 850/1; 851/1; 852/1; 853/1; 854/1; 855/1; 856/1; 857/1; 858/1; 859/1; 860/1; 861/1; 862/1; 863/1; 864/1; 865/1; 866/1; 867/1; 868/1; 869/1; 870/1; 871/1; 872/1; 873/1; 874/1; 875/1; 876/1; 877/1; 878/1; 879/1; 880/1; 881/1; 882/1; 883/1; 884/1; 885/1; 886/1; 887/1; 888/1; 889/1; 890/1; 891/1; 892/1; 893/1; 894/1; 895/1; 896/1; 897/1; 898/1; 899/1; 900/1; 901/1; 902/1; 903/1; 904/1; 905/1; 906/1; 907/1; 908/1; 909/1; 910/1; 911/1; 912/1; 913/1; 914/1; 915/1; 916/1; 917/1; 918/1; 919/1; 920/1; 921/1; 922/1; 923/1; 924/1; 925/1; 926/1; 927/1; 928/1; 929/1; 930/1; 931/1; 932/1; 933/1; 934/1; 935/1; 936/1; 937/1; 938/1; 939/1; 940/1; 941/1; 942/1; 943/1; 944/1; 945/1; 946/1; 947/1; 948/1; 949/1; 950/1; 951/1; 952/1; 953/1; 954/1; 955/1; 956/1; 957/1; 958/1; 959/1; 960/1; 961/1; 962/1; 963/1; 964/1; 965/1; 966/1; 967/1; 968/1; 969/1; 970/1; 971/1; 972/1; 973/1; 974/1; 975/1; 976/1; 977/1; 978/1; 979/1; 980/1; 981/1; 982/1; 983/1; 984/1; 985/1; 986/1; 987/1; 988/1; 989/1; 990/1; 991/1; 992/1; 993/1; 994/1; 995/1; 996/1; 997/1; 998/1; 999/1; 1000/1; 1001/1; 1002/1; 1003/1; 1004/1; 1005/1; 1006/1; 1007/1; 1008/1; 1009/1; 1010/1; 1011/1; 1012/1; 1013/1; 1014/1; 1015/1; 1016/1; 1017/1; 1018/1; 1019/1; 1020/1; 1021/1; 1022/1; 1023/1; 1024/1; 1025/1; 1026/1; 1027/1; 1028/1; 1029/1; 1030/1; 1031/1; 1032/1; 1033/1; 1034/1; 1035/1; 1036/1; 1037/1; 1038/1; 1039/1; 1040/1; 1041/1; 1042/1; 1043/1; 1044/1; 1045/1; 1046/1; 1047/1; 1048/1; 1049/1; 1050/1; 1051/1; 1052/1; 1053/1; 1054/1; 1055/1; 1056/1; 1057/1; 1058/1; 1059/1; 1060/1; 1061/1; 1062/1; 1063/1; 1064/1; 1065/1; 1066/1; 1067/1; 1068/1; 1069/1; 1070/1; 1071/1; 1072/1; 1073/1; 1074/1; 1075/1; 1076/1; 1077/1; 1078/1; 1079/1; 1080/1; 1081/1; 1082/1; 1083/1; 1084/1; 1085/1; 1086/1; 1087/1; 1088/1; 1089/1; 1090/1; 1091/1; 1092/1; 1093/1; 1094/1; 1095/1; 1096/1; 1097/1; 1098/1; 1099/1; 1100/1; 1101/1; 1102/1; 1103/1; 1104/1; 1105/1; 1106/1; 1107/1; 1108/1; 1109/1; 1110/1; 1111/1; 1112/1; 1113/1; 1114/1; 1115/1; 1116/1; 1117/1; 1118/1; 1119/1; 1120/1; 1
```

RESULT 5  
T13954  
MEGF6 protein - rat



Biochem. J. 295, 131-140, 1993  
A;Title: Identification of the predominant glycosaminoglycan-attachment site in soluble serine.  
A;Reference number: S38954; MUID:94029900; PMID:8216207  
A;Accession: S38954  
A;Molecule type: protein  
A;Residues: 475-491;X', 493-494 <GER>  
A;Note: the residue designated 'X' was determined to be a Ser with covalently bound chon  
R;Meininger, D.P.; Komives, E.A.  
submitted to the Brookhaven Protein Data Bank, September 1995  
A;Reference number: A67369; PDB:1ZQA  
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
R;Tulinsky, A.; Mathews, I.I.  
submitted to the Brookhaven Protein Data Bank, August 1994  
A;Reference number: A52804; PDB:1HLT  
A;Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442  
R;Hrabal, R.; Komives, E.A.; Ni, F.  
submitted to the Brookhaven Protein Data Bank, November 1995  
A;Reference number: A65583; PDB:1PGD  
A;Contents: annotation; conformation by (1)H-NMR, residues 427-444  
R;Hrabal, R.; Komives, E.A.; Ni, F.  
Protein Sci. 5, 195-203, 1996  
A;Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the  
A;Reference number: A58595; MUID:96276211; PMID:8745396  
A;Contents: annotation; conformation by (1)H-NMR  
C;Genetics:  
A;Gene: GDB:THBD  
A;Cross-references: GDB:119613; OMIM:188040  
A;Map position: 20p11.2-20p11.2  
A;Introns: #status absent  
C;Complex: homodimer, urinary form  
C;Function:  
A;Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activation  
A;Pathway: blood coagulation moderation  
A;Note: the membrane-bound form is located on the endothelium luminal surface of arterie  
A;Note: thrombin complexed with the membrane-bound form is subject to endocytosis  
C;Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C;Keywords: anticoagulant; beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coag  
e protein  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>  
F;19-573/Domain: extracellular #status predicted <EXT>  
F;19-486/Product: thrombomodulin, urinary form #status experimental <MAU>  
F;24-167/Domain: C-type lectin homology <LCH>  
F;177-199/Region: PEST sequence  
F;201-233/Region: PEST sequence  
F;245-280/Domain: EGF homology <EG1>  
F;288-323/Domain: EGF homology <EG2>  
F;329-362/Domain: EGF homology <EG3>  
F;369-404/Domain: EGF homology <EG4>  
F;408-439/Domain: EGF homology <EG5>  
F;445-480/Domain: EGF homology <EG6>  
F;485-513/Region: PEST sequence  
F;517-539/Domain: transmembrane #status predicted <TMN>  
F;540-575/Domain: intracellular #status predicted <INT>  
F;47,115,116,382,409/Binding site: carbonylate (Asn) (covalent) #status predicted  
F;174,225,411,504/Binding site: carbonylate (Thr) (covalent) #status predicted  
F;245-256,252-265,267-280,288-292,308,310-323,329-340,336-349,351-362,369-378,374-38  
F;334,498/Binding site: carbonylate (Ser) (covalent) #status predicted  
F;342/Modified site: erythro-beta-hydroxyasparagine (Asn) #status experimental  
F;490,492/Binding site: Chondroitin sulfate (Ser) (covalent) (partial) #status experimen

Query Match 100.0%; Score 2916; DB 1; Length 575;  
Best Local Similarity 100.0%; Pred. No. 5.1e-169;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVVLGALALAGLGFAPAPPGGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MGVVLGALALAGLGFAPAPPGGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60

QY 61 TVRSSVAADVISLLNGDGGVRRRLWLGQLPPGCGPKRLGRLGFGQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVISLLNGDGGVRRRLWLGQLPPGCGPKRLGRLGFGQWVTGDNNTSYS 120

QY 121 RWRDLNGAPLCGLPCVAVSAAEATVPSEPTEWEEQCEVKADGFLCEFHFPATCRPLAV 180  
DB 121 RWRDLNGAPLCGLPCVAVSAAEATVPSEPTEWEEQCEVKADGFLCEFHFPATCRPLAV 180

QY 181 EPGAAAAVSIYTGTFPAARGADFOALPYGSSAAVAPLGLQMLCTAPPAGAVOCHWAREAP 240  
DB 181 EPGAAAAVSIYTGTFPAARGADFOALPYGSSAAVAPLGLQMLCTAPPAGAVOCHWAREAP 240

QY 241 GAWDCSVENGCGEHCACNAIPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVNP 300

QY 301 DQPGSYSCMCEYGRLAADQHRCEVDVDDCILEPSPCQRCVNTQGGFECHECPNYDLVDG 360  
DB 301 DQPGSYSCMCEYGRLAADQHRCEVDVDDCILEPSPCQRCVNTQGGFECHECPNYDLVDG 360

QY 361 ECVPEVDPGFRANCEYQOCPLNOTSYLVCACGFAFIPHEPHRCOMFCNOTACPADCDPN 420  
DB 361 ECVPEVDPGFRANCEYQOCPLNOTSYLVCACGFAFIPHEPHRCOMFCNOTACPADCDPN 420

QY 421 TQASCEPGEYILDDGFICTDIDECENGCGFCGVCNHLPGTFECICGPDALSALVHRHIGTDC 480  
DB 421 TQASCEPGEYILDDGFICTDIDECENGCGFCGVCNHLPGTFECICGPDALSALVHRHIGTDC 480

QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPPAVGLVHSG 516  
DB 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 2  
A60501  
Thrombomodulin precursor - mouse  
N;Alternate names: fetomodulin  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Jul-1999  
C;Accession: S08488; A32001; A60501  
R;Dittman, W.A.; Majerus, P.W.  
Nucleic Acids Res. 17, 802, 1989  
A;Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted  
A;Reference number: S08488; MUID:89128454; PMID:2536925  
A;Accession: S08488  
A;Molecule type: mRNA  
A;Residues: 1-577 <DIT>  
A;Cross-references: EMBL:X14432; NID:g54781; PIDN:CAA32597.1; PID:g54782  
R;Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.  
J. Biol. Chem. 263, 15815-15822, 1988  
A;Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetat  
A;Reference number: A32001; MUID:89008498; PMID:2844823  
A;Accession: A32001  
A;Molecule type: mRNA  
A;Residues: 97-577 <D12>  
A;Cross-references: GB:J04060  
R;Imada, S.; Yamaguchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.  
Dev. Biol. 140, 113-122, 1990  
A;Title: Identification of fetomodulin, a surface marker protein of fetal development  
A;Reference number: A60501; MUID:90292331; PMID:2162790  
A;Accession: A60501  
A;Molecule type: protein  
A;Residues: 19-22;330-343;479-489;545-555;562-575 <IMA>  
C;Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofac  
C;Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C;Keywords: anticoagulant; membrane protein; phosphoprotein; receptor  
F;24-165/Domain: C-type lectin homology <LCH>  
F;244-329/Domain: EGF homology <EG1>  
F;287-322/Domain: EGF homology <EG2>  
F;328-361/Domain: EGF homology <EG3>  
F;368-403/Domain: EGF homology <EG4>  
F;407-438/Domain: EGF homology <EG5>  
F;444-479/Domain: EGF homology <EG6>

Query Match 65.1%; Score 1899; DB 2; Length 577;  
Best Local Similarity 65.5%; Pred. No. 1.3e-107;

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:22:58 ; Search time 16.5 seconds

(without alignments)  
3006.386 Million cell updates/sec

Title: US-09-509-994-1

Perfect score: 2916

Sequence: 1 MLGVLVGLALAGLGPAP.....PSPTPGSLTPPAVGLVHSG 516

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: p1r1.\*

2: p1r2.\*

3: p1r3.\*

4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2916	100.0	575	1 THHUB	thrombomodulin pre
2	1899	65.1	577	2 A60501	thrombomodulin pre
3	1092.5	37.5	356	2 A25918	thrombomodulin - b
4	416	14.3	1620	2 T27283	hypothetical prote
5	414	14.2	1574	2 T13954	MEGF6 protein - ra
6	373	12.8	2907	2 A52728	fibrillin-2 precu
7	367.5	12.6	1184	2 A55184	fibulin-2 precuso
8	367	12.6	2871	2 A55567	fibrillin 1 - bovi
9	365	12.5	2918	2 A54105	fibrillin-2 precu
10	363	12.4	3002	2 A47231	fibrillin-1 precu
11	361	12.4	2871	2 A55624	fibrillin-1 precu
12	359	12.3	741	2 T46488	hypothetical prote
13	357	12.2	1221	2 T49457	fibulin-2 precuso
14	331	11.4	589	2 T43210	fibulin-1D precurs
15	329.5	11.3	1964	2 T09059	notch4 - mouse
16	327.5	11.2	1712	2 A38261	masking protein pr
17	324.5	11.1	689	2 T42760	fibulin, splice fo
18	324.5	11.1	712	2 T42990	fibulin 1, splice
19	320.5	11.0	2321	2 T78549	notch3 protein - h
20	320	11.0	1394	2 A35626	transforming growt
21	314.5	10.8	3507	2 T34513	hypothetical prote
22	313	10.7	601	2 B36346	fibulin 1 precuso
23	313	10.7	683	2 C36346	fibulin 1 precuso
24	313	10.7	798	2 T22793	hypothetical prote
25	312.5	10.7	1820	2 A55494	latent transformin
26	308	10.6	685	2 T78040	fibulin, splice fo
27	308	10.6	705	2 S34968	fibulin, splice fo
28	303.5	10.4	1251	2 A57293	latent transformin
29	296	10.2	2531	2 A46019	Notch-1 protein -

30	286	9.8	2471	2 A49128	cell-fate determin
31	285	9.8	2352	2 T30201	Notch homolog prot
32	283.5	9.7	2437	2 S42612	transmembrane prot
33	280.5	9.6	2703	1 A24420	notch protein - fr
34	280	9.6	2318	2 S45306	notch 3 protein -
35	278.5	9.6	387	2 T38449	extracellular prote
36	277.5	9.5	511	2 T17298	hypothetical prote
37	277	9.5	2531	2 S18188	notch protein homo
38	276.5	9.5	493	2 JC5621	epidermal growth f
39	272	9.3	2555	2 A40043	notch protein homo
40	269.5	9.2	1217	1 EGM5MG	epidermal growth f
41	263	9.0	1203	2 A49175	Notch B protein -
42	259.5	8.9	2531	2 T31070	notch homolog - se
43	256.5	8.8	2524	2 A35844	Notch protein - Af
44	254.5	8.7	674	2 T55476	growth potentiatio
45	249.5	8.6	1133	1 EGR7	epidermal growth f

## ALIGNMENTS

### RESULT 1

#### THHUB

thrombomodulin precursor [validated] - human

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1988 #sequence\_revision 12-May-1995 #text\_Change 15-Sep-2000

C:Accession: A41442; A28307; A29680; A27073; JX0264; S38954

R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruy

J. Biochem. 103, 281-285, 1988

A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed ac

A:Reference number: A41442; MUID:88227901; PMID:2836377

A:Accession: A41442

A:Molecule type: DNA

A:Residues: 1-575 <SHI>

A:Cross-references: DBJ:D00210; NID:g220126; PIDN:BAA00149.1; PID:g220127

R:Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.

Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987

A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of the

A:Reference number: A28307; MUID:87317665; PMID:2819876

A:Accession: A28307

A:Molecule type: DNA; mRNA

A:Residues: 1-472, 'A', 474-575 <JAC>

A:Cross-references: GB:J03973; NID:g339658; PIDN:AAA61175.1; PID:g339659

R:Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioaka, J.; Maruyama, I.; Zushi, M.; Kaw

EMBO J. 6, 1891-1897, 1987

A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on end

A:Reference number: A29680; MUID:88004395; PMID:2820710

A:Accession: A29680

A:Molecule type: mRNA

A:Residues: 1-575 <SUZ>

A:Cross-references: GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:g736251

A:Experimental source: lung endothelium

A:Note: part of this sequence, including the amino end of the mature protein, were de

R:Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.

Biochemistry 26, 4350-4357, 1987

A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of

A:Reference number: A27073; MUID:88024950; PMID:2822087

A:Accession: A27073

A:Molecule type: mRNA

A:Residues: 1-472, 'A', 474-575 <WEN>

A:Cross-references: GB:M16552; NID:g339656; PIDN:AAB59508.1; PID:g339657

A:Experimental source: placenta

A:Note: parts of this sequence were determined by protein sequencing

R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.

J. Biochem. 113, 433-440, 1993

A:Title: Urinary thrombomodulin, its isolation and characterization.

A:Reference number: JX0264; MUID:93293792; PMID:8390445

A:Accession: JX0264

A:Molecule type: protein; mRNA

A:Residues: 19-472, 'A', 474-486 <YAM>

A:Experimental source: urine

A:Note: the urinary form appears to be identical with that circulating in plasma

R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.



```
xx Recombinant protein having thrombomodulin activity -
PT used in the therapeutic control of coagulation and the treatment
PT and prevent of thrombotic episodes
xx
PS Disclosure; ; p; English.
xx
CC Protein is encoded by plasmid p2.1 derived from human cell
CC line A549 known to express about 10000 molecules of thrombomodulin
CC per cell. Plasmid p2.1 showed a strong hybridisation signal with a
CC 60-mer bovine thrombomodulin probe. Thrombomodulins are used to
CC potentiate a patient's anticoagulant capacity.
CC See also AAN82027 and AAN82037.
xx
SQ Sequence 575 AA;

Query Match          99.7%; Score 2906; DB 9; Length 575;
Best Local Similarity 99.6%; Pred. No. 1.7e-161;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGVVLVGLALAGLGFPAEPPGPGSQVHCDFALYPGPATFLNASQICDGLRGLHM 60
DB 1 MGVVLVGLALAGLGFPAEPPGPGSQVHCDFALYPGPATFLNASQICDGLRGLHM 60

QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPCCGDPKRLGPIRGFOWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPCCGDPKRLGPIRGFOWVTGDNNTSYS 120

QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVSEPIWEEQQCEVKADGFLCEFFHPATCRPLAV 180
DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVSEPIWEEQQCEVKADGFLCEFFHPATCRPLAV 180

QY 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240

QY 241 GAWDCSVENGCGEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300
DB 241 GAWDCSVENGCGEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300

QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360
DB 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360

QY 361 ECVPEVDFCFRANCYQCPQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
DB 361 ECVPEVDFCFRANCYQCPQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TQASCECEPGYLLDDGFICTDIDECENGFCGVCYCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 TQASCECEPGYLLDDGFICTDIDECENGFCGVCYCHNLPGTFECICGPDLSALVRHIGTDC 480

QY 481 DSGKYDGDGDSGGEPPSPPTGSLTTPPAVGLVHSG 516
DB 481 DSGKYDGDGDSGGEPPSPPTGSLTTPPAVGLVHSG 516
```

Search completed: December 16, 2002, 17:24:24  
Job time : 37.5 secs

Db 301 DQGSYSCHMETGYRLAADOHRCEVDVDCILLEPSPQRCVNTQGGFECHECYPNYDLVDG 360  
Qy 361 ECVEPVDPCFRANCEYQCQLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
Db 361 ECVEPVAPCFRANCEYQCQLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
Qy 421 TQASCEPCEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Db 421 TQASCEPCEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Qy 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 14  
AAR10617  
ID AAR10617 standard; Protein; 515 AA.  
XX  
AC  
DT 17-APR-1991 (first entry)  
XX  
XX Soluble thrombomodulin deriv.  
XX Thrombosis; anticoagulant; Protein C.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..16  
FT /label= signal peptide  
FT /note= "for low mol. wt deriv.; optionally present"  
FT Peptide 1..18  
FT /label= signal peptide  
FT /note= "for high mol wt. deriv; optionally present"  
XX EP412841-A.  
XX  
PD 13-FEB-1991.  
XX  
PF 10-AUG-1990; 90EP-0308826.  
XX  
PR 05-FEB-1990; 90US-0474870.  
PR 11-AUG-1989; 89US-0393617.  
XX  
PA (ELIL ) ELI LILLY & CO.  
XX  
PI Bang NU, Grinnell BW, Hoskins JA, Moore RE, Parkinson JF;  
XX  
XX WPI; 1991-045960/07.  
XX N-PSDB; AAQ10435.  
XX  
PT New soluble thrombomodulin derivatives - produced by recombinant  
PT DNA techniques for use as anticoagulants and in treatment of  
PT thrombosis.  
XX  
PS Claim 2; Page 45; 81pp; English.  
XX  
CC The sequence is that of a sol. thrombomodulin deriv. comprising the  
CC signal peptide (optional), N-terminal, epidermal growth factor  
CC homology region and the Ser/Thr-rich region, but lacking the trans-  
CC membrane and cytoplasmic domains. The sequence was deduced DNA  
CC obtd. from clone GHTW3A isolated from a human chromosome 20 library,  
CC subcloned into pUC19 to give pGHTW3A. A Ppu MI fragment of this  
CC vector was ligated with a linker to construct pUC18TM, which was  
CC treated with BsmI to delete about 500 bp from the 3' end of the  
CC gene. ( The deletion occurs at the Ser/Thr-rich/transmembrane domain  
CC junction, effectively cleaving the transmembrane and cytoplasmic  
CC domains.) The plasmid was recircularised to give pUC18TMD for use  
CC in prodn of the recombinant sol. deriv. The protein reacts with  
CC thrombin to activate the Protein C anticoagulant pathway, inhibit  
CC thrombin, and can be used to treat or prevent thrombotic disorders.

XX SQ Sequence 515 AA;  
Query Match 99.7%; Score 2906; DB 12; Length 515;  
Best Local Similarity 99.8%; Pred. No. 1.5e-161;  
Matches 514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MGVLVLCALALAGLGFAPAPAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MGVLVLCALALAGLGFAPAPAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWICLQLPCCGPKRLGPGFQWVTDGNTNTSY 120  
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWICLQLPCCGPKRLGPGFQWVTDGNTNTSY 120  
Qy 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEQQCEVKADGFCEHFHFPATCRPLAV 180  
Db 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEQQCEVKADGFCEHFHFPATCRPLAV 180  
Qy 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGQLMCTAPPGAVQGHWAREAP 240  
Db 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGQLMCTAPPGAVQGHWAREAP 240  
Qy 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300  
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300  
Qy 301 DQGSYSCHMETGYRLAADOHRCEVDVDCILLEPSPQRCVNTQGGFECHECYPNYDLVDG 360  
Db 301 DQGSYSCHMETGYRLAADOHRCEVDVDCILLEPSPQRCVNTQGGFECHECYPNYDLVDG 360  
Qy 361 ECVEPVDPCFRANCEYQCQLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
Db 361 ECVEPVDPCFRANCEYQCQLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
Qy 421 TQASCEPCEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Db 421 TQASCEPCEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Qy 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 515  
Db 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 515

RESULT 15  
AAP82070  
ID AAP82070 standard; protein; 575 AA.  
XX  
AC AAP82070;  
XX  
DT 19-OCT-1990 (first entry)  
XX  
DE Human thrombomodulin encoded by plasmid p2.1.  
XX  
XX Thrombomodulin activity; Protein C; anticoagulant;  
KW epidermal Growth Factor (EGF) domains.  
XX  
OS synthetic.  
XX  
XX WO8809811-A.  
XX  
PD 15-DEC-1988.  
XX  
PF 09-JUN-1988; 88WO-DK00089.  
XX  
PR 12-JUN-1987; 87DK-0002990.  
XX  
PA (NOVO ) NOVO INDUSTRI A/S.  
XX  
PI Nexo BA, Esper B;  
XX  
XX WPI; 1988-368626/51.  
DR N-PSDB; AAP82026.

XX 28-JUL-2000 (first entry)  
 DT Human thrombomodulin TMD protein #2.  
 XX Human; thrombomodulin; vasculitis; protein C; thrombin.  
 XX Homo sapiens.  
 OS  
 XX JP2000053582-A.  
 PN  
 XX 22-FEB-2000.  
 PD  
 XX 06-AUG-1998; 98JP-0222688.  
 PF  
 XX 06-AUG-1998; 98JP-0222688.  
 PR  
 XX (ASAH ) ASAH KASEI KOGYO KK.  
 PA  
 XX WPI; 2000-353249/31.  
 DR  
 XX N-PSDB; AAA10028.  
 DR  
 XX Treating agent for vasculitis contains peptide which promotes  
 PT activation of protein C by thrombin -  
 PT  
 XX Claim 4; Page 13-14; 18pp; Japanese.  
 PS  
 XX This sequence represents a human thrombomodulin protein. The invention  
 CC relates to a treating agent for vasculitis containing a peptide which  
 CC promotes activation of protein C by thrombin.  
 CC  
 SQ Sequence 516 AA;

Query Match 99.9%; Score 2912; DB 21; Length 516;  
 Best Local Similarity 99.8%; Pred. No. 6.8e-162;  
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGVLVLGALAGLGFPPAPAEPPQGGSCQVEHDCFCALYPGPAPFLNASQICDGLRGLHM 60  
 DB 1 MGVLVLGALAGLGFPPAPAEPPQGGSCQVEHDCFCALYPGPAPFLNASQICDGLRGLHM 60  
 QY 61 TVRSSVAADVISLLNGDGGVRRRLWTGLQPPCGDPPKRLGRLGFWQWTDGNTSYS 120  
 DB 61 TVRSSVAADVISLLNGDGGVRRRLWTGLQPPCGDPPKRLGRLGFWQWTDGNTSYS 120  
 QY 121 RWARLDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180  
 DB 121 RWARLDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
 DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGEACNATPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
 DB 241 GAWDCSVENGCGEACNATPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDCCIPEPSPQRCVNTQGGFCHCYPNYDLDVG 360  
 DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDCCIPEPSPQRCVNTQGGFCHCYPNYDLDVG 360  
 QY 361 ECVEPVDFCFRANCEYQQPLNQTSLYLCVCAEGFAPIPHEHRCQFNCQACPADCDPN 420  
 DB 361 ECVEPVDFCFRANCEYQQPLNQTSLYLCVCAEGFAPIPHEHRCQFNCQACPADCDPN 420  
 QY 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCHNLPGTPECICGPDALVRHIGTDC 480  
 DB 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCHNLPGTPECICGPDALVRHIGTDC 480  
 QY 481 DSGKVDGDSGSGEPPPTGSTLTPPAVGLVHSG 516  
 DB 481 DSGKVDGDSGSGEPPPTGSTLTPPAVGLVHSG 516

RESULT 13  
 AAR22017  
 ID AAR22017 standard; Protein; 516 AA.  
 XX  
 XX AAR22017;  
 XX 03-JUL-1992 (first entry)  
 DT  
 XX Human thrombomodulin (1-516) with Asp367 substituted by Ala.  
 DE  
 XX Mutant; thrombin binding site; blood clotting; Tmml mutator.  
 KW  
 XX Homo sapiens.  
 OS  
 XX EP474273-A.  
 PN  
 XX 11-MAR-1992.  
 PD  
 XX 05-AUG-1991; 91EP-0202009.  
 PF  
 XX 03-AUG-1990; 90JP-0204978.  
 PR  
 XX (ASAH ) ASAH KASEI KOGYO.  
 PA  
 XX Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;  
 PI  
 XX WPI; 1992-081820/11.  
 DR  
 XX New polypeptide inhibits blood coagulation and platelet  
 CC aggregation - promotes thrombin catalysed protein C activation  
 CC for treating myocardial infarction, thrombosis, embolism, etc.  
 CC  
 SQ Example 1; Page 18; 112pp; English.

Plasmid M13TMD1 (see AAR22016) encoding the first 516 N-terminal  
 amino acids of human thrombomodulin was used as a template for  
 site-directed mutagenesis using the "Tmml mutator" to produce  
 plasmid M13TMD1. In the mutant thrombomodulin encoded by the  
 plasmid, the wild-type Asp residue at position 367 is substituted  
 by an Ala residue. This mutant, truncated thrombomodulin has low  
 activity in the activation of protein C, c.f. high activity for the  
 similarly truncated thrombomodulin with Asp 367.  
 See AAR22013-R22022 and AAQ25073.  
 SQ Sequence 516 AA;  
 Query Match 99.7%; Score 2908; DB 13; Length 516;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-161;  
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGVLVLGALAGLGFPPAPAEPPQGGSCQVEHDCFCALYPGPAPFLNASQICDGLRGLHM 60  
 DB 1 MGVLVLGALAGLGFPPAPAEPPQGGSCQVEHDCFCALYPGPAPFLNASQICDGLRGLHM 60  
 QY 61 TVRSSVAADVISLLNGDGGVRRRLWTGLQPPCGDPPKRLGRLGFWQWTDGNTSYS 120  
 DB 61 TVRSSVAADVISLLNGDGGVRRRLWTGLQPPCGDPPKRLGRLGFWQWTDGNTSYS 120  
 QY 121 RWARLDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180  
 DB 121 RWARLDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
 DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGEACNATPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
 DB 241 GAWDCSVENGCGEACNATPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDCCIPEPSPQRCVNTQGGFCHCYPNYDLDVG 360  
 DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDCCIPEPSPQRCVNTQGGFCHCYPNYDLDVG 360

Matches	515;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
1	MLGVVLGALAGLAGLGGFPAPAEPPQGSQCVEHDCFALYPGPATFLNASQICDGLRGHLM	60							
1	MLGVVLGALAGLAGLGGFPAPAEPPQGSQCVEHDCFALYPGPATFLNASQICDGLRGHLM	60							
61	TVRSVAADVISILLNGDGGVGRRRLWIGLQLPPEGCGDKRILGPRGFOWVTGDNNTSYS	120							
61	TVRSVAADVISILLNGDGGVGRRRLWIGLQLPPEGCGDKRILGPRGFOWVTGDNNTSYS	120							
121	RWARLDLNGAPLCGPLCVASAAEATVPSEPIWEEOQCEVKADGFLCEHFHPATCRPLAV	180							
121	RWARLDLNGAPLCGPLCVASAAEATVPSEPIWEEOQCEVKADGFLCEHFHPATCRPLAV	180							
181	EPGAAAAAIVTITGTTPFAARGADFQALPVGSSAAVAPLGQLMQTMCTAPCAVOGHWAREAP	240							
181	EPGAAAAAIVTITGTTPFAARGADFQALPVGSSAAVAPLGQLMQTMCTAPCAVOGHWAREAP	240							
241	GAWDCSVENGGECHACNAIPAIPRCQCPAGAALQADGRSCTASATQSCNDLCEHFVCVNP	300							
241	GAWDCSVENGGECHACNAIPAIPRCQCPAGAALQADGRSCTASATQSCNDLCEHFVCVNP	300							
301	DQPGSYSCMCETGYRLAADQHRCEDVDCCIILESPCPCQRVCNTQGGEFCHCYPTNYDLVDG	360							
301	DQPGSYSCMCETGYRLAADQHRCEDVDCCIILESPCPCQRVCNTQGGEFCHCYPTNYDLVDG	360							
361	ECVEPDPFRANCEYOQPLNOTSYLYCVAEGFATPIPEPHRCOMFCNOTACPADC DN	420							
361	ECVEPDPFRANCEYOQPLNOTSYLYCVAEGFATPIPEPHRCOMFCNOTACPADC DN	420							
421	TQASCEPEGYILLDDGFICTIDDECENGGFCSGVCHNLPGTFECICGPDALSALVRHIGTDC	480							
421	TQASCEPEGYILLDDGFICTIDDECENGGFCSGVCHNLPGTFECICGPDALSALVRHIGTDC	480							
481	DSGKVDGSDSGSGPPPSPTFGSTLTLPAPVGLVHS	516							
481	DSGKVDGSDSGSGPPPSPTFGSTLTLPAPVGLVHS	516							

RESULT 11	
RAY09348	
ID	AA09348 standard; Protein; 516 AA.
XX	
XX	AA09348;
XX	
XX	08-JUL-1999 (first entry)
DT	
DT	Human thrombomodulin SEQ ID NO:2.
XX	
XX	Human; thrombomodulin; aqueous parenteral solution; storage;
KW	distribution; acute coronary syndrome; thrombosis; embolism;
KW	diabetes.
KW	
XX	
XX	Homo sapiens.
OS	
XX	W09918994-A1.
XX	
XX	22-APR-1999.
PD	
PD	
XX	
XX	13-OCT-1998; 98WO-JP04609.
PF	
XX	
XX	11-NOV-1997; 97JP-0308523.
PR	
PR	15-OCT-1997; 97JP-0281659.
XX	
XX	(ASAH ) ASAH I KASEI KOGYO KK.
PPA	
XX	
XX	Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;
PI	
XX	
DR	WPI: 1999-277444/23.
DR	N-PSDB: AAX55880.
XX	
XX	Stable aqueous parenteral thrombomodulin solution - comprising
PPT	buffer and surfactant, useful for treating acute coronary syndrome
PPT	

thrombosis, embolism, and diabetes

Claim 7; Page 90-92; 97pp; Japanese.

The present invention describes a method for maintaining the quality of an aqueous parenteral solution of thrombomodulin comprising buffer and surfactant aseptically filled in a case or syringe. Maintaining the quality of an aqueous, parenteral thrombomodulin solution is characterised in that the solution: (a) comprises soluble thrombomodulin, a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically filled into: (i) a case or (ii) a syringe without any empty space; and (c) is kept in liquid form in storage and distribution and not frozen or freeze dried. Thrombomodulin is used to treat acute coronary syndrome (such as myocardial infarction, unstable angina and coronary artery blockage), thrombosis (e.g. cerebral, vascular and peripheral blood vessel thrombosis), embolism, peripheral blood vessel disorders (e.g. Raynauds disease), arteriosclerosis, vasculitis (e.g. due to systemic lupus erythematosus or Barrett's syndrome), multiple organ failure, disseminated intravascular coagulation, transient ischaemia, diabetes, liver veno-occlusive diseases and deep vein thrombosis. The composition is stable for a long period of time and can be stored and distributed in ready to use form avoiding the problems of dissolution and accuracy when preparing on demand. The present sequence represents human thrombomodulin.

Sequence 516 AA:

Query Match	99.9%	Score	2912;	DB	20;	Length	516;
Best Local Similarity	99.98;	Pred.	No. 6.8e-162;				
Matches	515;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps	0;						
Qy	1	MLGVVLGALALAGLGPAPAEPOPGGSGQCVHEHDCFALYPCGPATFLNASQICDGLRGHLM	60				
Db	1	MLGVVLGALALAGLGPAPAEPOPGGSGQCVHEHDCFALYPCGPATFLNASQICDGLRGHLM	60				
Qy	61	TVRSSVAADVISLLINGDGGVRRRLIGLQLPGCGDGPKRGLGRGFQWYTGDNNTSYS	120				
Db	61	TVRSSVAADVISLLINGDGGVRRRLIGLQLPGCGDGPKRGLGRGFQWYTGDNNTSYS	120				
Qy	121	RWARLDLNGAPLCPCLVAVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFHPATCRPLAV	180				
Db	121	RWARLDLNGAPLCPCLVAVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFHPATCRPLAV	180				
Qy	181	EPGAAAAVSITYTPTFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAQVGHWAREAP	240				
Db	181	EPGAAAAVSITYTPTFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAQVGHWAREAP	240				
Qy	241	GAWDCSVENGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP	300				
Db	241	GAWDCSVENGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP	300				
Qy	301	DQPGSYCMCETGYRLAADQHRCEVDVDCILEPSPCORCVNTGGFFCHCYPNYDLVDG	360				
Db	301	DQPGSYCMCETGYRLAADQHRCEVDVDCILEPSPCORCVNTGGFFCHCYPNYDLVDG	360				
Qy	361	ECVEPVPDCFRANCEYOCQPLNQSYLVCVAGGAPIPHEPHRCQMFNCQATCAPDCPN	420				
Db	361	ECVEPVPDCFRANCEYOCQPLNQSYLVCVAGGAPIPHEPHRCQMFNCQATCAPDCPN	420				
Qy	421	TOASCEPEGYILDDGFICTDIDECENGCGFCGVCNHLPGTFECICGPDLSALVRHIGTDC	480				
Db	421	TOASCEPEGYILDDGFICTDIDECENGCGFCGVCNHLPGTFECICGPDLSALVRHIGTDC	480				
Qy	481	DSGKVDGSDSGSGBPSPPTPGSLTTPPAVGLVHSG	516				
Db	481	DSGKVDGSDSGSGBPSPPTPGSLTTPPAVGLVHSG	516				

RESULT 12  
AAV83935  
ID AAV83935 standard; Protein; 516 AA.  
XX  
XX AC AAV83935;

FT Domain 19..244  
 FT /label= N-terminal\_domain  
 FT 245..480  
 FT /note= "6 EGF-like domains"  
 FT Modified-site 481..515  
 FT /note= "O-linked glycosylation site"  
 FT 516..539  
 FT /label= Transmembrane\_domain  
 FT 540..575  
 FT /label= Cytoplasmic\_domain  
 PN W0200198352-A2.  
 XX 27-DEC-2001.  
 XX 20-JUN-2001; 2001WO-US19590.  
 XX 21-JUN-2000; 2000US-213678P.  
 PR 12-JUN-2001; 2001US-0880484.  
 XX (SCHD ) SCHERING AG.  
 XX Light D, Morser MJ, Nagashima M;  
 PI N-PSDB; AAD28182.  
 DR WPI; 2002-154622/20.  
 DR N-PSDB; AAD28182.  
 XX Thrombomodulin analog with greater ability to potentiate  
 PT thrombin-mediated activation of protein C and lesser ability to  
 PT potentiate thrombin-mediated activation of thrombin-activatable  
 PT fibrinolysis inhibitor for treating thrombotic disease -  
 XX Claim 1; Fig 4; 53pp; English.  
 XX The invention relates to the design, production and use of thrombomodulin  
 CC (TM) analogues that have the ability to enhance the thrombin-mediated  
 CC activation of protein C which have a significantly reduced ability to  
 CC promote activation of thrombin-activatable fibrinolysis inhibitor (TAFI).  
 CC TM analogue is useful for treating thrombotic disease in human, by  
 CC administering it or a polynucleotide encoding it to the patient and  
 CC expressing the analogue in vivo. Composition comprising TM analogue is  
 CC useful for treating a thrombotic disease or condition in mammals. TM  
 CC analogue is useful in antithrombotic therapy. TM analogue is also useful  
 CC for treating diseases in which thrombus formation plays a significant  
 CC etiological role including myocardial infarction, disseminated  
 CC intravascular coagulation (DIC), deep vein thrombosis, septic shock,  
 CC pulmonary embolism, acute respiratory distress syndrome, unstable angina  
 CC and other arterial or venous occlusive conditions. TM analogue is also  
 CC useful for treating various circulatory disorders including coronary or  
 CC pulmonary embolism, stroke and systemic coagulation disorders including  
 CC DIC which is often associated with septicemia, certain cancers and  
 CC toxemia of pregnancy. The present sequence is human full-length native  
 CC thrombomodulin.  
 XX Sequence 575 AA;  
 SQ Query Match 100.0%; Score 2916; DB 23; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-162;  
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLCVLVIGALALAGLGFPAEPQPGSQVHCDFALYPGPATFLNASQICDGLRGHLM 60  
 DB 1 MLCVLVIGALALAGLGFPAEPQPGSQVHCDFALYPGPATFLNASQICDGLRGHLM 60  
 QY 61 TVRSSVAADVISILLNGDGVGRRRLWIGLQLPPGCGDPRKLGRLGFWQVNTGNTSYS 120  
 DB 61 TVRSSVAADVISILLNGDGVGRRRLWIGLQLPPGCGDPRKLGRLGFWQVNTGNTSYS 120  
 QY 121 RWARLDNGLAPLCGLPCVAVSAEAIVPSEPITWEEQCEVKADGFLCEHFFPACRPLAV 180  
 DB 121 RWARLDNGLAPLCGLPCVAVSAEAIVPSEPITWEEQCEVKADGFLCEHFFPACRPLAV 180  
 QY 181 EPGAAAAVSIYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240

Db 181 EPGAAAAVSIYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
 QY 241 GAWDCSVNGGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 Db 241 GAWDCSVNGGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCEYGLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHCYPNYDLVDG 360  
 Db 301 DQPGSYSCMCEYGLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHCYPNYDLVDG 360  
 QY 361 ECVEPVDFCFRANCEYQCQPLNOTSYLCVCAEGFAPIDPEHRCOMFCNQACPADCDPN 420  
 Db 361 ECVEPVDFCFRANCEYQCQPLNOTSYLCVCAEGFAPIDPEHRCOMFCNQACPADCDPN 420  
 QY 421 TQASCECEGYLLDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALVRHIGTDC 480  
 Db 421 TQASCECEGYLLDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALVRHIGTDC 480  
 QY 481 DSGKVDGDSGSGEPSPPTFGSTLTPPAVGLVHSG 516  
 Db 481 DSGKVDGDSGSGEPSPPTFGSTLTPPAVGLVHSG 516  
 RESULT 10  
 AAR22018  
 ID AAR22018 standard; Protein; 516 AA.  
 XX AAR22018;  
 AC AAR22018;  
 XX 03-JUL-1992 (first entry)  
 DT Human thrombomodulin (1-516) with Asp367 substituted by Glu.  
 XX DE  
 XX KW Mutant; thrombin binding site; blood clotting; TMm2 mutator.  
 XX OS Homo sapiens.  
 XX PN EP474273-A.  
 XX PD 11-MAR-1992.  
 XX PF 05-AUG-1991; 91EP-0202009.  
 XX PR 03-AUG-1990; 90JP-0204978.  
 XX (ASAH ) ASAH KASEI KOGYO.  
 PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;  
 DR WPI; 1992-081820/11.  
 XX New polypeptide inhibits blood coagulation and platelet  
 PT aggregation - promotes thrombin catalysed protein C activation  
 PT for treating myocardial infarction, thrombosis, embolism, etc.  
 XX Example 1; Page 18; 112pp; English.  
 XX Plasmid M13PMD1 (see AAR22016) encoding the first 516 N-terminal  
 CC amino acids of human thrombomodulin was used as a template for  
 CC site-directed mutagenesis using the "TMm2 mutator" to produce  
 CC plasmid M13TM2. In the mutant thrombomodulin encoded by the  
 CC plasmid, the wild-type Asp residue at position 367 is substituted  
 CC by an Glu residue. The activity of this mutant, truncated  
 CC thrombomodulin (i.e. D123Glu) in the activation of protein C was  
 CC compared to that of similarly truncated thrombomodulin with Asp  
 CC at position 367 (i.e. D123Asp). The activity of D123Glu was as high  
 CC as 1.2 times that of D123Asp. See AAR22013-R22022 and AAQ25074.  
 XX Sequence 516 AA;  
 SQ Query Match 99.9%; Score 2912; DB 13; Length 516;  
 Best Local Similarity 99.8%; Pred. No. 6.8e-162;



|||||  
Db 361 ECVEPDPCEFRANCEYQCQLNTSYLCVCAEGFAPIPHEPHKQMCNQTACPADCDPN 420  
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 516  
RESULT 8  
AAE23026  
ID AAE23026 standard; Protein; 575 AA.  
XX  
AC AAE23026;  
XX  
DT 21-AUG-2002 (first entry)  
XX  
DE Human thrombomodulin protein.  
XX  
KW Human; thrombomodulin; analogue; neurologic trauma; vulneryary;  
KW spinal cord injury; neurological damage; therapy; tranquilliser;  
KW neuroprotective.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT Peptide 1..18  
FT Protein /label= Signal\_peptide  
FT /note= "Human mature thrombomodulin"  
FT Domain 1..226  
FT /note= "N-terminal domain"  
FT Domain 227..462  
FT /note= "6 EGF-like domain"  
FT Modified-site 463..497  
FT /note= "O-linked glycosylation site"  
FT Domain 498..521  
FT /note= "Transmembrane domain"  
FT Domain 522..557  
FT /note= "Cytoplasmic domain"  
XX  
PN WO200217953-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 31-AUG-2001; 2001WO-US41930.  
XX  
PR 31-AUG-2000; 2000US-229714P.  
XX 23-AUG-2001; 2001US-0938405.  
XX  
PA (SCHD ) SCHERING AG.  
XX  
PI Festoff BW, Morser MJ;  
XX  
DR WPI; 2002-415551/44.  
XX N-PSDB; AAD36901.  
XX  
PT Treating neurological damage resulting from spinal cord injury in  
PT humans by administering a soluble, recombinant thrombomodulin analog  
PT which is resistant to oxidation -  
XX  
XX  
PS Claim 1; Fig 1; 31pp; English.  
XX  
CC The invention relates to a method of using thrombomodulin analogues  
CC in the treatment of the neurologic trauma associated with spinal  
CC cord injury in mammals. The method involves administering a oxidation  
CC resistant, soluble, recombinant thrombomodulin analogue, where the  
CC methionine at position 388 has been replaced with a leucine, and the  
CC analogue is numbered in accordance with native thrombomodulin. The  
CC method and thrombomodulin analogue are useful for treating

CC neurological damage resulting from spinal cord injury in human.  
CC The present sequence is human thrombomodulin protein.  
XX  
XX Sequence 575 AA;  
Query Match 100.0%; Score 2916; DB 23; Length 575;  
Best Local Similarity 100.0%; Pred. No. 4.4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLGVLVLGALALAGLGFPPAPAEQPGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60  
Db 1 MLGVLVLGALALAGLGFPPAPAEQPGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60  
QY 61 TVRSVAADVISLLINGDGGVRRRLWTGLQIPGCGDPKRLGLRGQWGTGDNNTSYS 120  
Db 61 TVRSVAADVISLLINGDGGVRRRLWTGLQIPGCGDPKRLGLRGQWGTGDNNTSYS 120  
QY 121 RWRDLNGAPLCPLCVAVSAAEATVPSEPIWEQQCEVKADGFLCEFHFPATCRPLAV 180  
Db 121 RWRDLNGAPLCPLCVAVSAAEATVPSEPIWEQQCEVKADGFLCEFHFPATCRPLAV 180  
QY 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCAGALQADGRSCTASATQSCNDLCEHFCVPPNP 300  
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCAGALQADGRSCTASATQSCNDLCEHFCVPPNP 300  
QY 301 DPGSYSCMCEGYRLAADOHRCEVDVDCILEPSPORCVNTQGGFECHECHYPNTDLVDG 360  
Db 301 DPGSYSCMCEGYRLAADOHRCEVDVDCILEPSPORCVNTQGGFECHECHYPNTDLVDG 360  
QY 361 ECVEPDPCEFRANCEYQCQLNTSYLCVCAEGFAPIPHEPHKQMCNQTACPADCDPN 420  
Db 361 ECVEPDPCEFRANCEYQCQLNTSYLCVCAEGFAPIPHEPHKQMCNQTACPADCDPN 420  
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 516  
RESULT 9  
AAE17521  
ID AAE17521 standard; Protein; 575 AA.  
XX  
AC AAE17521;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Human full-length native thrombomodulin (TM).  
XX  
KW Human; thrombomodulin; TM analogue; protein C; thrombotic disease;  
KW antithrombotic therapy; thrombin-activatable fibrinolysis inhibitor;  
KW TAFI; myocardial infarction; disseminated intravascular coagulation;  
KW DIC; deep vein thrombosis; septic shock; pulmonary embolism; angina;  
KW acute respiratory distress syndrome; cancer; toxemia; septicemia;  
KW circulatory disorder; coronary embolism; pulmonary embolism; stroke;  
KW systemic coagulation disorder; immunosuppressive; cerebroprotective;  
KW pregnancy; anticoagulant; thrombolytic; venous occlusive condition;  
KW antibacterial; cardiant.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..18  
FT Protein /label= Signal\_peptide  
FT /note= "Human mature thrombomodulin"

```

PR 30-JUL-1991; 91JP-0189984.
XX (ASAH ) ASahi CHEM IND CO LTD.
XX WPI; 1993-299652/38.
XX Novel polypeptide obtd. by culturing transformed fungus - having
PT blood coagulation preventing, platelet aggregation preventing and
PT thrombolytic activities
XX Disclosure; Fig 55; 65pp; Japanese.
XX Homo sapiens.
XX WO9322447-A.
XX 11-NOV-1993.
XX 30-APR-1993; 93WO-JP00578.
XX 01-MAY-1992; 92JP-0112903.
XX (ASAH ) ASahi CHEM IND CO LTD.
XX (ASAH ) ASahi KASEI KOGYO KK.
XX Kondo S, Toma K, Zushi M;
XX WPI; 1993-368806/46.
XX Peptide with anticoagulant and platelet aggregation inhibitor
PT activity - which promotes protein C activation by thrombin and is
PT useful in treating coagulation disorders e.g. thrombosis
XX Disclosure; Fig 1; 84pp; Japanese.
XX New peptides (see AAR50069) are inhibitors of the blood coagulation
CC and platelet aggregation activities of thrombin and promote the
CC protein-C activation effect of thrombin. They can be produced
CC efficiently in pure form by culture of appropriate transformants,
CC and are useful in treatment of circulatory disorders such as
CC myocardial infarction, thrombosis, embolism, telangiectasis,
CC arteriosclerosis obliterans, disseminated intravascular
CC coagulation, angina pectoris, gestosis and transient ischaemic
CC attack.
XX SQ Sequence 575 AA;
Query Match 100.0%; Score 2916; DB 14; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLGVLVGLALAGLGPAPAEPPQPGSQCVHEHDFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MLGVLVGLALAGLGPAPAEPPQPGSQCVHEHDFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQPPCGDPRKRLGRLGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQPPCGDPRKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180
Db 121 RWARLDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180
QY 181 EPGAAAANVSITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240
Db 181 EPGAAAANVSITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQPCQRCVNTQGGFECYCHYNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQPCQRCVNTQGGFECYCHYNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCQPLNNTSYLCVCAEGFAPIPHEPHRCQFCNQTACPDNDPN 420
Db 361 ECVEPVDPCFRANCEYQCQPLNNTSYLCVCAEGFAPIPHEPHRCQFCNQTACPDNDPN 420
QY 421 TQASCECPGEGYILDDGICTDDCEBNGFCSCGVCHNLPGTFFCICGPOSALYRHIGTDC 480
Db 421 TQASCECPGEGYILDDGICTDDCEBNGFCSCGVCHNLPGTFFCICGPOSALYRHIGTDC 480
QY 481 DSGKVDGSGSGGEPPTPGSTLTPPAVGLVHSG 516
Db 481 DSGKVDGSGSGGEPPTPGSTLTPPAVGLVHSG 516
RESULT 7
AAR43031
ID AAR43031 standard; protein; 575 AA.
XX
AC AAR43031;
XX
DT 16-MAY-1994 (first entry)
XX

```

```

DE Human thrombomodulin.
XX Anticoagulant; platelet aggregation inhibitor;
XX protein C; activation; thrombin; thrombomodulin;
KW coagulation disorder; thrombosis; myocardial infarction;
KW embolism; telangiectasis; arteriosclerosis obliterans;
KW disseminated intravascular coagulation; DIC; angina pectoris;
KW gestosis; transient ischaemic attack.
XX OS
XX Homo sapiens.
XX WO9322447-A.
XX 11-NOV-1993.
XX 30-APR-1993; 93WO-JP00578.
XX 01-MAY-1992; 92JP-0112903.
XX (ASAH ) ASahi CHEM IND CO LTD.
XX (ASAH ) ASahi KASEI KOGYO KK.
XX Kondo S, Toma K, Zushi M;
XX WPI; 1993-368806/46.
XX Peptide with anticoagulant and platelet aggregation inhibitor
PT activity - which promotes protein C activation by thrombin and is
PT useful in treating coagulation disorders e.g. thrombosis
XX Disclosure; Fig 1; 84pp; Japanese.
XX New peptides (see AAR50069) are inhibitors of the blood coagulation
CC and platelet aggregation activities of thrombin and promote the
CC protein-C activation effect of thrombin. They can be produced
CC efficiently in pure form by culture of appropriate transformants,
CC and are useful in treatment of circulatory disorders such as
CC myocardial infarction, thrombosis, embolism, telangiectasis,
CC arteriosclerosis obliterans, disseminated intravascular
CC coagulation, angina pectoris, gestosis and transient ischaemic
CC attack.
XX SQ Sequence 575 AA;
Query Match 100.0%; Score 2916; DB 14; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLGVLVGLALAGLGPAPAEPPQPGSQCVHEHDFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MLGVLVGLALAGLGPAPAEPPQPGSQCVHEHDFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQPPCGDPRKRLGRLGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQPPCGDPRKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180
Db 121 RWARLDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180
QY 181 EPGAAAANVSITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240
Db 181 EPGAAAANVSITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQPCQRCVNTQGGFECYCHYNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQPCQRCVNTQGGFECYCHYNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCQPLNNTSYLCVCAEGFAPIPHEPHRCQFCNQTACPDNDPN 420

```



QY 421 TQASCEPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDPSALVRHIGTDC 480  
 Db 421 TQASCEPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDPSALVRHIGTDC 480  
 QY 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
 Db 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 3  
 AAY83934  
 ID AAY83934 standard; Protein; 516 AA.  
 XX  
 AC AAY83934;  
 XX  
 DT 28-JUL-2000 (first entry)  
 XX Human thrombomodulin TMD protein.  
 DE Human; thrombomodulin; vasculitis; protein C; thrombin.  
 KW Homo sapiens.  
 OS JP2000053582-A.  
 PN 22-FEB-2000.  
 PD 06-AUG-1998; 98JP-0222688.  
 PF 06-AUG-1998; 98JP-0222688.  
 PR (ASAH ) ASAH KASEI KOGYO KK.  
 PA WPI; 2000-353249/31.  
 DR N-PSDB; AAA10027.  
 XX  
 PT Treating agent for vasculitis contains peptide which promotes  
 activation of protein C by thrombin -  
 XX  
 PS Claim 4; Page 10-12; 18pp; Japanese.  
 CC This sequence represents a human thrombomodulin protein. The invention  
 relates to a treating agent for vasculitis containing a peptide which  
 promotes activation of protein C by thrombin.  
 XX  
 SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 4e-162;  
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVLVLGALALAGLGFAPAEPOPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 Db 1 MGVLVLGALALAGLGFAPAEPOPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQLPPCGDPRKRLGFWTGDNNNTSYS 120  
 Db 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQLPPCGDPRKRLGFWTGDNNNTSYS 120  
 QY 121 RWARLDNGAPLCGLPVAVSAAEATVPSEPIWEQOCEVADGFLCFEHPFATCRPLAV 180  
 Db 121 RWARLDNGAPLCGLPVAVSAAEATVPSEPIWEQOCEVADGFLCFEHPFATCRPLAV 180  
 QY 181 EFGAAAAVSYITGTFPAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVGHWAREAP 240  
 Db 181 EFGAAAAVSYITGTFPAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVGHWAREAP 240  
 QY 241 GAWDCSVENGCGEACNAIPGAPROCCPAGAAQADGSRCTASATQSCNDLCEHFCVNP 300  
 Db 241 GAWDCSVENGCGEACNAIPGAPROCCPAGAAQADGSRCTASATQSCNDLCEHFCVNP 300  
 QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILPSPQPCVNTQGGFECHECHYPNYDLVDG 360  
 Db 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILPSPQPCVNTQGGFECHECHYPNYDLVDG 360

Db 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILPSPQPCVNTQGGFECHECHYPNYDLVDG 360  
 QY 361 ECVEPVDPCFRANCEYOCQPLNQTSLYCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
 Db 361 ECVEPVDPCFRANCEYOCQPLNQTSLYCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
 QY 421 TQASCEPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDPSALVRHIGTDC 480  
 Db 421 TQASCEPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDPSALVRHIGTDC 480  
 QY 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
 Db 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 4  
 AAY69529  
 ID AAY69529 standard; Protein; 516 AA.  
 XX  
 AC AAY69529;  
 XX  
 DT 10-APR-2000 (first entry)  
 XX Human thrombomodulin variant, SEQ ID NO:3.  
 DE Thrombomodulin; TM; soluble; affinity purification; cation exchange;  
 KW anticoagulant; thrombolytic.  
 OS Homo sapiens.  
 OS JP11341990-A.  
 PN 14-DEC-1999.  
 PD 23-MAR-1999; 99JP-0077518.  
 PF 30-MAR-1998; 98JP-0084389.  
 PR (ASAH ) ASAH KASEI KOGYO KK.  
 PA WPI; 2000-101696/09.  
 DR N-PSDB; AAZ55965.  
 XX Preparation of highly pure soluble thrombomodulin - used as an  
 antilob coagulation agent and a thrombolytic agent  
 PT  
 PS Claim 9; Page 30-32; 36pp; Japanese.  
 CC The invention relates to a novel method for the preparation of highly  
 pure soluble thrombomodulin (TM) containing substantially no serum-  
 derived or antibody-derived substance. The method comprises isolating  
 soluble TM from unpurified serum via affinity purification using an  
 anti-TM antibody. The soluble TM is then treated with a cation  
 exchanger at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to  
 4, and the fraction containing the soluble TM is isolated. The soluble  
 TM can be used as a blood anticoagulation agent and a thrombolytic agent.  
 CC This sequence represents a human thrombomodulin variant, designated  
 SEQ ID NO:3.  
 XX  
 SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 4e-162;  
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVLVLGALALAGLGFAPAEPOPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 Db 1 MGVLVLGALALAGLGFAPAEPOPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQLPPCGDPRKRLGFWTGDNNNTSYS 120  
 Db 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQLPPCGDPRKRLGFWTGDNNNTSYS 120

XX Plasmid M13mp19TMJ3 (see AAR22013) was used as template for  
CC site-directed mutagenesis. A 177bp fragment was deleted using the  
CC "TMD1 deleter" oligonucleotide to give plasmid M13TMD1 which encodes  
CC the first 516 N-terminal amino acids of human thrombomodulin. Plasmid  
CC M13TMD1 was completely digested with HindIII and BamHI and a TMD1  
CC -fragment of ca. 1700bp was isolated. The fragment was ligated to  
CC HindIII- and BglII-cut plasmid pSV2-dhfr to give the recombinant  
CC plasmid pSV2TMD1. See AAR22014-R22022 and AAQ25072.  
XX  
SQ Sequence 516 AA;  
  
Query Match 100.0%; Score 2916; DB 13; Length 516;  
Best Local Similarity 100.0%; Pred. No. 4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MGVVLGALALAGLGPAPAEPPGSGQCVHDCFCALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MGVVLGALALAGLGPAPAEPPGSGQCVHDCFCALYPGPATFLNASQICDGLRGHLM 60  
  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPCGDPRKRLGPFQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPCGDPRKRLGPFQWVTGDNNTSYS 120  
  
QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
  
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
  
QY 241 GAWDCSVENGCGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
  
QY 301 DQPGSYSCMCEYGYLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
DB 301 DQPGSYSCMCEYGYLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
  
QY 361 ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCPN 420  
DB 361 ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCPN 420  
  
QY 421 TOASCEPEGYILDGFICTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 480  
DB 421 TOASCEPEGYILDGFICTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 480  
  
QY 481 DSGKVDGSDSGGEPSPPTPGSTILTPPAVGLVHSG 516  
DB 481 DSGKVDGSDSGGEPSPPTPGSTILTPPAVGLVHSG 516  
  
RESULT 2  
AA09347  
ID AA09347 standard; Protein: 516 AA.  
XX  
AC AA09347;  
XX  
DT 08-JUL-1999 (first entry)  
XX  
DE Human thrombomodulin SEQ ID NO:1.  
XX  
KW Human; thrombomodulin; aqueous parenteral solution; storage;  
KW distribution; acute coronary syndrome; thrombosis; embolism;  
KW diabetes.  
XX  
OS Homo sapiens.  
XX  
PN WO9918994-A1.  
XX  
PD 22-APR-1999.  
XX

PF 13-OCT-1998; 98WO-JP04609.  
XX  
PR 11-NOV-1997; 97JP-0308523.  
PR 15-OCT-1997; 97JP-0281659.  
XX  
PA (ASAH ) ASahi KASEI KOGYO KK.  
XX  
PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;  
XX  
DR WPI; 1999-277444/23.  
DR N-PSDB; AAX55879.  
XX  
PT Stable aqueous parenteral thrombomodulin solution - comprising  
PT buffer and surfactant, useful for treating acute coronary syndrome,  
PT thrombosis, embolism, and diabetes  
XX  
PS Claim 6; Page 87-89; 97pp; Japanese.  
XX  
CC The present invention describes a method for maintaining the quality of  
CC an aqueous parenteral solution of thrombomodulin comprising buffer and  
CC surfactant aseptically filled in a case or syringe. Maintaining the  
CC quality of an aqueous, parenteral thrombomodulin solution is  
CC characterised in that the solution: (a) comprises soluble thrombomodulin,  
CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically  
CC filled into: (i) a case or (ii) a syringe without any empty space; and  
CC (c) is kept in liquid form in storage and distribution and not frozen or  
CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome  
CC (such as myocardial infarction, unstable angina and coronary artery  
CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood  
CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.  
CC Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic  
CC lupus erythematosus or Barrett's syndrome), multiple organ failure,  
CC disseminated intravascular coagulation, transient ischaemia, diabetes,  
CC liver veno-occlusive diseases and deep vein thrombosis. The composition  
CC is stable for a long period of time and can be stored and distributed in  
CC ready to use form avoiding the problems of dissolution and accuracy when  
CC preparing on demand. The present sequence represents human  
CC thrombomodulin.  
XX  
SQ Sequence 516 AA;  
  
Query Match 100.0%; Score 2916; DB 20; Length 516;  
Best Local Similarity 100.0%; Pred. No. 4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MGVVLGALALAGLGPAPAEPPGSGQCVHDCFCALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MGVVLGALALAGLGPAPAEPPGSGQCVHDCFCALYPGPATFLNASQICDGLRGHLM 60  
  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPCGDPRKRLGPFQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPCGDPRKRLGPFQWVTGDNNTSYS 120  
  
QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
  
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
  
QY 241 GAWDCSVENGCGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
  
QY 301 DQPGSYSCMCEYGYLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
DB 301 DQPGSYSCMCEYGYLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
  
QY 361 ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCPN 420  
DB 361 ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCPN 420

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:19:38 ; Search time 34.5 Seconds  
(without alignments)  
1992.966 Million cell updates/sec

Title: US-09-509-994-1  
Perfect score: 2916  
Sequence: 1 MLGVLGALALAGLFPAP.....PSPTPGSTLPPAVGLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues 908470  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2916	100.0	516	13	Truncated human th
2	2916	100.0	516	20	Human thrombomodul
3	2916	100.0	516	21	Human thrombomodul
4	2916	100.0	516	21	Human thrombomodul
5	2916	100.0	575	12	Human thrombomodul
6	2916	100.0	575	14	Thrombomodulin. H
7	2916	100.0	575	14	Human thrombomodul
8	2916	100.0	575	23	Human thrombomodul
9	2916	100.0	575	23	Human full-length
10	2912	99.9	516	13	Human thrombomodul

11	2912	99.9	516	20	AA109348	Human thrombomodul
12	2912	99.9	516	21	AA183935	Human thrombomodul
13	2908	99.7	516	13	AA122017	Human thrombomodul
14	2906	99.7	515	12	AA10617	Soluble thrombomod
15	2906	99.7	575	9	AA182070	Human thrombomodul
16	2904	99.6	575	14	AA131572	Human thrombomodul
17	2899	99.4	575	13	AA120639	Human urinary thro
18	2875	98.6	575	20	AA173970	Human thrombomodul
19	2858	98.0	575	13	AA122189	Sequence of thromb
20	2830	97.1	498	16	AA184185	Thrombomodulin TME
21	2830	97.1	498	18	AA101600	Human thrombomodul
22	2827	96.9	557	23	AA123028	Human thrombomodul
23	2827	96.9	557	23	AA123031	Human thrombomodul
24	2827	96.9	557	23	AA117525	Human thrombomodul
25	2827	96.9	557	23	AA117528	Human thrombomodul
26	2826	96.9	557	23	AA117523	Human thrombomodul
27	2825	96.9	498	21	AA167402	Novel sugar chain-
28	2825	96.9	557	23	AA117596	Human thrombomodul
29	2824	96.8	497	19	AA169520	rsfM protein SEQ I
30	2824	96.8	557	23	AA117530	Human thrombomodul
31	2824	96.8	557	23	AA117598	Human thrombomodul
32	2823	96.8	557	23	AA123029	Human thrombomodul
33	2823	96.8	557	23	AA117526	Human thrombomodul
34	2822	96.8	557	23	AA123030	Human thrombomodul
35	2822	96.8	557	23	AA117527	Human thrombomodul
36	2822	96.8	557	23	AA117529	Human thrombomodul
37	2822	96.8	557	23	AA117593	Human thrombomodul
38	2822	96.8	557	23	AA117595	Human thrombomodul
39	2822	96.8	557	23	AA117597	Human thrombomodul
40	2821	96.7	557	23	AA117591	Human thrombomodul
41	2821	96.7	557	23	AA117592	Human thrombomodul
42	2821	96.7	557	23	AA117594	Human thrombomodul
43	2820	96.7	497	17	AA194607	Human recombinant
44	2820	96.7	557	23	AA117524	Human thrombomodul
45	2815	96.5	554	23	AA123027	Human thrombomodul

ALIGNMENTS

RESULT 1

AA122016

ID AA122016 standard; Protein; 516 AA.

XX AA122016;

AC AC

XX 03-JUL-1992 (first entry)

XX DE

XX DE Truncated human thrombomodulin encoded by plasmid pSV2TMD1.

XX KW

XX KW Thrombin binding site; blood clotting; TMD1 deleter.

XX OS

XX OS Homo sapiens.

XX PN

XX PN EP474273-A.

XX PD

XX PD 11-MAR-1992.

XX PF

XX PF 05-AUG-1991; 91EP-0202009.

XX PR

XX PR 03-AUG-1990; 90JP-0204978.

XX PA

XX PA (ASAH ) ASAH KASEI KOGYO.

XX PI

XX PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;

XX DR

XX DR WPI; 1992-081820/11.

XX PT

XX PT New polypeptide inhibits blood coagulation and platelet

XX PT aggregation - promotes thrombin catalysed protein C activation

XX PT for treating myocardial infarction, thrombosis, embolism, etc.

XX PS

XX PS Example 1; Page 17; 112pp; English.

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:26:23 ; Search time 8.5 Seconds  
(without alignments)  
1011.502 Million cell updates/sec

Title: US-09-509-994-1  
Perfect score: 2916  
Sequence: 1 MLGVLVGLALAGLGFAP.....PSPTPGSTLPPAVGLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105981 seqs, 16662342 residues  
Total number of hits satisfying chosen parameters: 105981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2916	100.0	575	10	US-09-938-405-2
2	581	19.9	652	10	US-09-789-919-96
3	543	18.6	644	10	US-09-789-919-62
4	543	18.6	644	10	US-09-789-919-73
5	355	12.2	776	9	US-10-000-512-8
6	355	12.2	959	9	US-10-000-512-10
7	352.5	12.1	877	10	US-09-764-898-200
8	350.5	12.0	915	9	US-09-905-291A-34
9	350.5	12.0	915	10	US-09-909-320-34
10	350.5	12.0	915	10	US-09-909-088B-34
11	333	11.4	999	10	US-09-747-371-2
12	328.5	11.3	1009	10	US-09-898-570-16
13	327.5	11.2	974	10	US-09-898-570-14
14	326	11.2	800	10	US-09-764-853-800
15	326	11.2	800	10	US-09-764-898-272
16	323	11.1	534	10	US-09-804-156-14
17	323	11.1	534	10	US-09-946-633-6
18	315.5	10.8	845	10	US-09-898-570-12
19	315	10.8	997	10	US-09-747-371-3
20	311	10.7	433	10	US-09-925-302-691
21	308	10.6	448	9	US-10-041-016-2
22	308	10.6	448	10	US-09-083-002-2
23	308	10.6	448	10	US-09-275-805-2
24	308	10.6	448	10	US-09-836-561-1
25	302	10.4	448	9	US-10-066-500-15
26	302	10.4	448	12	US-10-052-586-408
27	288.5	9.9	493	9	US-09-974-298-52
28	288	9.9	735	10	US-09-898-570-10
29	278.5	9.6	387	10	US-09-836-561-5
30	277	9.5	368	10	US-09-275-805-9
31	271	9.3	2444	10	US-09-944-849-2
32	268	9.2	1055	10	US-09-855-722-2
33	267	9.2	1212	10	US-09-855-722-3
34	267	9.2	1238	10	US-09-855-722-5
35	267	9.2	1238	10	US-09-944-849-4
36	266.5	9.1	509	9	US-09-905-291A-315
37	266.5	9.1	509	10	US-09-909-320-315
38	266.5	9.1	509	10	US-09-909-088B-315
39	266.5	9.1	509	12	US-10-052-586-52
40	251.5	8.6	816	10	US-09-976-165-37
41	241	8.3	886	10	US-09-992-647-11
42	234.5	8.0	362	10	US-09-764-898-226
43	234	8.0	722	10	US-09-908-322-12
44	233.5	8.0	1404	10	US-09-944-849-8
45	230.5	7.9	1246	10	US-09-919-497-85

ALIGNMENTS

RESULT 1  
US-09-938-405-2  
; Sequence 2, Application US/09938405  
; Patent No. US20020111296A1  
; GENERAL INFORMATION:  
; APPLICANT: Restoff, Barry W.  
; APPLICANT: Morser, Michael J.  
; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Injury  
; FILE REFERENCE: 51960AUSMI  
; CURRENT APPLICATION NUMBER: US/09/938,405  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: 60/229,714  
; PRIOR FILING DATE: 2000-08-31  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-938-405-2

Query Match	100.0%	Score	2916;	DB	10;	Length	575;
Best Local Similarity	100.0%	Pred	No. 2,4e-174;				
Matches	516;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MLGVLVGLALAGLGFAPPAEPQPGSQCVHDCFCALYPGPATFLNASQICDGLRGLHM	60				
Db	1	MLGVLVGLALAGLGFAPPAEPQPGSQCVHDCFCALYPGPATFLNASQICDGLRGLHM	60				
QY	61	TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGGDPRRLGRLGFWGTGDNNTSYS	120				
Db	61	TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGGDPRRLGRLGFWGTGDNNTSYS	120				
QY	121	RWARDLNGAPLGGPLCVASAAEATVPSEPIWEQOCEVKADGFLCEFFHFPATCRPLAV	180				
Db	121	RWARDLNGAPLGGPLCVASAAEATVPSEPIWEQOCEVKADGFLCEFFHFPATCRPLAV	180				
QY	181	EPGAAAAAIVSYTGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP	240				
Db	181	EPGAAAAAIVSYTGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP	240				
QY	241	GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP	300				



```
Db 241 GAWDCSVENGCGEACNAIPGRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Qy 301 DOPGYSYCMCTGYRLAQRHCEVDVDCILEPSPQRCQVNTQGGFCHCYPNYDLVDG 360
Db 301 DOPGYSYCMCTGYRLAQRHCEVDVDCILEPSPQRCQVNTQGGFCHCYPNYDLVDG 360
Qy 361 ECVEPVPDPCFRANCEYQCPNLQNTSYLCVCAEGFAPIPHEPHRCQMFNCNTACADCDPN 420
Db 361 ECVEPVPDPCFRANCEYQCPNLQNTSYLCVCAEGFAPIPHEPHRCQMFNCNTACADCDPN 420
Qy 421 TOASCEPEGVILDDGFTCTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 TOASCEPEGVILDDGFTCTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Qy 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVLVHSG 516
Db 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVLVHSG 516

RESULT 2
US-09-789-919-96
; Sequence 96, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 96
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-789-919-96
```

```
Query Match 19.9%; Score 581; DB 10; Length 652;
Best Local Similarity 31.6%; Pred. No. 2.8e-29;
Matches 180; Conservative 56; Mismatches 210; Indels 124; Gaps 32;

Qy 2 LGVLVIGALALA--GLGFAPAPBPQSGSOVCHDCFALYPGPATFLNASQICDGLRHGL 59
Db 5 MGLLLLLLLLLTOPGAGTGADTE-----AVVCGTACTYTAHSGKLSAAEAQNHCCNQNGNL 60

Qy 60 MTRVSSVAAD-----VISLLNGDGGVGR--RLWIGLQLPPG--CGDPKRLGRLGFOWVT 112
Db 61 ATVKSKEEAHVQVRLAQLLRREALTARMSKFWIGLQREKGLCLDPSL--PLKGFSSWG 118

Qy 113 GDNNTSYRWARDLNGAPLCGLPVA--VSAEAATVPSE--PIWEEQOC-----EYKAD 163
Db 119 GGEDTPYSNNHKELRNSC--ISKRCVSLDLQSLPLNPKWSEGPCSGPGSPGSNIE 176

Qy 164 GFICEHFHPRATPLAV-EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLQL 222
Db 177 GFVCKFSFGKMCRLALGGPG-----QVTTYTPFTQTTSSLEAVPASAANVA----- 224

Qy 223 MC-TAPPGAVQGHW---AREAPGAWD-----CSVENGGCEHAC--NAIPGAP 263
Db 225 -CGEGDKDETQSHYFLCKEKAPDFWNGSSGPLCVSPKYCNFNNGCHODCFEGGDSF 283

Qy 264 RCQCPAGAAQADGRSCTASATQSCNDLCE--HFCVNPDPQGSYSQMCCTGYRLAADOH 321
Db 284 LCGRCPGRRLLDLVTC-ASRNPSSSPCRGGATCVLGP-HGKNYTCRCPPGQQLDSSQL 341

Qy 322 RCEDVDVDCILEPSPQRCQVNTQGGFCHCYPNYDLVDGECVBPDPFCFRANCEYQCPQL 381
Db 342 DCVDVDEC--QDSPCAQECVNTPGFRCECWGY-----EPGGP-----GEGACQDV 386

Qy 382 NQTSYLCVCAEGFAPIPHEPHRCQMFNCNTACADCDPNQAS--CECPGYIL--DDGF 437
```

```
Db 387 DE-----CALGRSP-----CAQGC-TNTDGSFHCSEEGYVLAGEGT 423
Qy 438 ICTDIDEC--ENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKVV-----D 486
Db 424 QCQDVDECVGPGGLCDLSLFCNTQGSFHCGLPGVWLAPN-GVSCMTGMPVSLGPPSGPPD 482
Qy 487 GGDGSGSGE-----PPSPPTPGSTLTTPPA 509
Db 483 EEDKGEKEGSTVPRAATASPTRGPEGTPKA 512

RESULT 3
US-09-789-919-62
; Sequence 62, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Katerl
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-789-919-62

Query Match 18.6%; Score 543; DB 10; Length 644;
Best Local Similarity 30.0%; Pred. No. 6.3e-27;
Matches 167; Conservative 55; Mismatches 211; Indels 124; Gaps 27;

Qy 5 LVIGALALAGLGFAPAPBPQSGSOVCHDCFALYPGPATFLNASQICDGLRHGLMTVRS 64
Db 9 LLLGLLQOPWAG--AAADSQ--AVVCGTACTYTAHSGKLSAAEAQNHCCNQNGNLATVKS 64

Qy 65 SVAA-----DVISLLNGDGGVGR--RLWIGLQLPPGCGDPKRLGRLGFOWVTGDNNTS 118
Db 65 EEEARHVQQAQLTQLLTKAPLEAKWGFWIGLQREKGNCTYHDL-PMRGFSWVGGEDETA 123

Qy 119 YSRWARDLNGAPLCGLPVAVSAEAATVPSE--PIWEEQOCV-----KADGFICEPHF 171
Db 124 YSNWYKASKSSCIFKRCVSLILDLSLTPHPSHLPKWHESPCTGTPPEAPGNSIEGLCKNF 183

Qy 172 PATCRPLAV-EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQ-----LM 223
Db 184 KGMCRPLALGGPG-----RVTTYTPFTQATTSSLEAVPFAVANVA-CGDEAKSETHYFL 236

Qy 224 CT-APPGAVQGHWAREAP---GAWDCSVENGCGEAC--NAIPGAPRCQCPAGAAQADG 277
Db 237 CNEKTPGIF--HWGSSGPLCVSPKFGCSFNNGCQQDCFEFGDGSFRCGPGPFRLLDDL 294

Qy 278 RSCTASATQSCNDLCEHFCVNPDPQPG-----SYSCMCTGYRLAADOHRCEDVD 327
Db 295 VTCA-----RNPCSSNPCTGGMCHSVPLSENYTCRCPSGQQLDSSQVHCVDDID 344

Qy 328 DCILEPSPQRCQVNTQGGFCHCYPNYDLVDG---ECVEPVPDPCFRANCEYQCPQLNQT 384
Db 345 EC--QDSPCAQCDVNTLGSFHCCEWGYQ-PSGPKKEACEDVDCAANSP----- 392

Qy 385 SYLCVCAEGFAPIPHEPHRCQMFNCNTACADCDPNQAS--CECPGYIL--DDGFICT 440
Db 393 -----CAQGC-----NDGSGFYCKEYIVSGEDSTQCE 423

Qy 441 DIDECE--GGFGSGVCHNLPGTFECICGPDLSALVRHIGTDCDSG-----KVD 486
Db 424 DIDECDARNPCDLSLFCNTDGSFRCCPGWELAPN-GVFCSGTGVFSELPAAPPOKED 482

Qy 487 GGDGSGSGPPSPPTPGS 503
```













OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (201)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (687)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:19:38 ; Search time 34.5 Seconds  
(without alignments)  
1992.966 Million cell updates/sec

Title: US-09-509-994-2

Perfect score: 2916  
Sequence: 1 MLGVILVGLALAGLFPAP.....PSPPTGSLTPPAVLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2916	100.0	516	20	AAV09348
2	2916	100.0	516	21	AAV83935
3	2912	99.9	516	13	AAE22016
4	2912	99.9	516	20	AAV09347
5	2912	99.9	516	21	AAV83934
6	2912	99.9	516	21	AAV69529
7	2912	99.9	516	12	AAV11534
8	2912	99.9	575	14	AAV41806
9	2912	99.9	575	14	AAV43031
10	2912	99.9	575	23	AAE23026

11	2912	99.9	575	23	AAE17521	Human full-length
12	2910	99.8	515	12	AAE10617	Soluble thrombomod
13	2910	99.8	575	9	AAV82070	Human thrombomodul
14	2908	99.7	516	13	AAE22018	Human thrombomodul
15	2908	99.7	575	14	AAE31572	Human thrombomodul
16	2904	99.6	516	13	AAE22017	Human thrombomodul
17	2903	99.6	575	13	AAE20639	Human urinary thro
18	2879	98.7	575	20	AAW73970	Human thrombomodul
19	2862	98.1	575	13	AAE22189	Sequence of thromb
20	2826	96.9	498	16	AAE84185	Human derived thro
21	2826	96.9	498	18	AAW01600	Thrombomodulin TME
22	2826	96.9	498	21	AAV67402	Novel sugar chain-
23	2824	96.8	497	17	AAE94607	Human recombinant
24	2823	96.8	557	23	AAE23028	Human thrombomodul
25	2823	96.8	557	23	AAE23031	Human thrombomodul
26	2823	96.8	557	23	AAE17525	Human thrombomodul
27	2823	96.8	557	23	AAE17528	Human thrombomodul
28	2822	96.8	557	23	AAE17523	Human thrombomodul
29	2821	96.7	557	23	AAE17596	Human thrombomodul
30	2820	96.7	497	19	AAW69520	rSTM protein SEQ I
31	2820	96.7	557	23	AAE17530	Human thrombomodul
32	2820	96.7	557	23	AAE17598	Human thrombomodul
33	2819	96.7	557	23	AAE23029	Human thrombomodul
34	2819	96.7	557	23	AAE17526	Human thrombomodul
35	2818	96.6	557	23	AAE23030	Human thrombomodul
36	2818	96.6	557	23	AAE17527	Human thrombomodul
37	2818	96.6	557	23	AAE17529	Human thrombomodul
38	2818	96.6	557	23	AAE17593	Human thrombomodul
39	2818	96.6	557	23	AAE17595	Human thrombomodul
40	2818	96.6	557	23	AAE17597	Human thrombomodul
41	2817	96.6	557	23	AAE17591	Human thrombomodul
42	2817	96.6	557	23	AAE17592	Human thrombomodul
43	2817	96.6	557	23	AAE17594	Human thrombomodul
44	2816	96.6	557	23	AAE17524	Human thrombomodul
45	2811	96.4	554	23	AAE23027	Human thrombomodul

ALIGNMENTS

RESULT 1	
AAV09348	
ID	AAV09348 standard; Protein; 516 AA.
XX	AAV09348;
AC	AAV09348;
XX	
DT	08-JUL-1999 (first entry)
XX	
DE	Human thrombomodulin SEQ ID NO:2.
XX	
KW	Human; thrombomodulin; aqueous parenteral solution; storage;
KW	distribution; acute coronary syndrome; thrombosis; embolism;
KW	diabetes.
XX	
OS	Homo sapiens.
XX	
PN	WO9918994-Al.
XX	
PD	22-APR-1999.
XX	
PF	13-OCT-1998; 98WO-JP04609.
XX	
PR	11-NOV-1997; 97JP-0308523.
PR	15-OCT-1997; 97JP-0281659.
XX	
PA	(ASAH ) ASahi Kasei Kogyo KK.
XX	
PI	Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;
XX	
DR	WPI: 1999-277444/23.
DR	N-PSDB; AAX55880.
XX	
PT	Stable aqueous parenteral thrombomodulin solution - comprising

PT buffer and surfactant, useful for treating acute coronary syndrome,  
 PT thrombosis, embolism, and diabetes  
 XX  
 PS  
 XX  
 XX Claim 7; Page 90-92; 97pp; Japanese.  
 CC The present invention describes a method for maintaining the quality of  
 CC an aqueous parenteral solution of thrombomodulin comprising buffer and  
 CC surfactant aseptically filled in a case or syringe. Maintaining the  
 CC quality of an aqueous, parenteral thrombomodulin solution is  
 CC characterised in that the solution: (a) comprises soluble thrombomodulin,  
 CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically  
 CC filled into: (i) a case or (ii) a syringe without any empty space; and  
 CC (c) is kept in liquid form in storage and distribution and not frozen or  
 CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome  
 CC (such as myocardial infarction, unstable angina and coronary artery  
 CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood  
 CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.  
 CC Raynauds disease), arteriosclerosis, vasculitis (e.g. due to systemic  
 CC lupus erythematosus or harrett's syndrome), multiple organ failure,  
 CC disseminated intravascular coagulation, transient ischaemia, diabetes,  
 CC liver veno-occlusive diseases and deep vein thrombosis. The composition  
 CC is stable for a long period of time and can be stored and distributed in  
 CC ready to use form avoiding the problems of dissolution and accuracy when  
 CC preparing on demand. The present sequence represents human  
 CC thrombomodulin.

XX  
 SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 20; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-163;  
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 DB 1 MLGVLVGLALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFGQVWTGDNNTSYS 120  
 DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFGQVWTGDNNTSYS 120  
 QY 121 RWARLDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFFPATCRPLAV 180  
 DB 121 RWARLDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFFPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
 DB 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
 DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
 QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
 DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
 QY 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
 DB 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
 QY 481 DSGKVDGSGSGEGPPPTPGSTLTTPPAVGLVHSG 516  
 DB 481 DSGKVDGSGSGEGPPPTPGSTLTTPPAVGLVHSG 516

RESULT 2

AA83935

ID AA83935 standard; Protein; 516 AA.

XX

AC  
 XX

28-JUL-2000 (first entry)

Human thrombomodulin TMD protein #2.

Human; thrombomodulin; vasculitis; protein C; thrombin.

Homo sapiens.

JP2000053582-A.

22-FEB-2000.

06-AUG-1998; 98JP-0222688.

06-AUG-1998; 98JP-0222688.

(ASAH ) ASAH KASEI KOGYO KK.

WPI: 2000-353249/31.

DR N-PSDB; AAL10028.

Treating agent for vasculitis contains peptide which promotes  
 PT activation of protein C by thrombin -

PS Claim 4; Page 13-14; 18pp; Japanese.

XX This sequence represents a human thrombomodulin protein. The invention  
 CC relates to a treating agent for vasculitis containing a peptide which  
 CC promotes activation of protein C by thrombin.

SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-163;  
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 DB 1 MLGVLVGLALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFGQVWTGDNNTSYS 120  
 DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFGQVWTGDNNTSYS 120  
 QY 121 RWARLDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFFPATCRPLAV 180  
 DB 121 RWARLDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFFPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
 DB 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
 DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
 QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
 DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
 QY 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
 DB 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
 QY 481 DSGKVDGSGSGEGPPPTPGSTLTTPPAVGLVHSG 516  
 DB 481 DSGKVDGSGSGEGPPPTPGSTLTTPPAVGLVHSG 516





```

XX SQ Sequence 516 AA;
Query Match 99.9%; Score 2912; DB 20; Length 516;
Best Local Similarity 99.8%; Pred. No. 1.3e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVLAGALAGLGFPAEPAPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLAGALAGLGFPAEPAPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGPLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGPLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCBHFVCVNP 300
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCBHFVCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQPCVNTQGGFECCHCYPNYDLVDG 360
DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQPCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCDPN 420
DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCDPN 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALARRHIGTDC 480
DB 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALARRHIGTDC 480
QY 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 5
AAY83934
ID AAY83934 standard; Protein; 516 AA.
XX AC AAY83934;
XX DT 28-JUL-2000 (first entry)
XX DE Human thrombomodulin TMD protein.
XX KW Human; thrombomodulin; vasculitis; protein C; thrombin.
XX OS Homo sapiens.
XX PN JP2000053582-A.
XX PD 22-FEB-2000.
XX FF 06-AUG-1998; 98JP-0222688.
XX PR 06-AUG-1998; 98JP-0222688.
XX PA (ASAH ) ASAH KASEI KOGYO KK.
XX DR WPI; 2000-353249/31.
XX DR N-PSDB; AAA10027.
XX PT Treating agent for vasculitis contains peptide which promotes
activation of protein C by thrombin.

```

```

XX PS Claim 4; Page 10-12; 18pp; Japanese.
XX CC This sequence represents a human thrombomodulin protein. The invention
CC relates to a treating agent for vasculitis containing a peptide which
CC promotes activation of protein C by thrombin.
XX SQ Sequence 516 AA;
Query Match 99.9%; Score 2912; DB 21; Length 516;
Best Local Similarity 99.8%; Pred. No. 1.3e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVLAGALAGLGFPAEPAPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLAGALAGLGFPAEPAPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGPLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGPLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCBHFVCVNP 300
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCBHFVCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQPCVNTQGGFECCHCYPNYDLVDG 360
DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQPCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCDPN 420
DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCDPN 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALARRHIGTDC 480
DB 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALARRHIGTDC 480
QY 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 6
AAY69529
ID AAY69529 standard; Protein; 516 AA.
XX AC AAY69529;
XX DT 10-APR-2000 (first entry)
XX DE Human thrombomodulin variant, SEQ ID NO:3.
XX KW Thrombomodulin; TM; soluble; affinity purification; cation exchange;
XX OS Homo sapiens.
XX PN JP11341990-A.
XX PD 14-DEC-1999.
XX PR 23-MAR-1999; 99JP-0077518.
XX PR 30-MAR-1998; 98JP-0084389.
XX

```

PA (ASAH ) ASahi KASEI KOGYO KK.  
XX WPI; 2000-101696/09.  
DR N-PSDB; AA255965.  
XX  
PT Preparation of highly pure soluble thrombomodulin - used as an  
XX antithrombotic agent and a thrombolytic agent  
XX  
PS Claim 9; Page 30-32; 36pp; Japanese.  
XX  
XX The invention relates to a novel method for the preparation of highly  
CC pure soluble thrombomodulin (TM) containing substantially no serum-  
CC derived or antibody-derived substance. The method comprises isolating  
CC soluble TM from unpurified serum via affinity purification using an  
CC anti-TM antibody. The soluble TM is then treated with a cation  
CC exchanger at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to  
CC 4, and the fraction containing the soluble TM is isolated. The soluble  
CC TM can be used as a blood anticoagulation agent and a thrombolytic agent.  
CC This sequence represents a human thrombomodulin variant, designated  
CC SEQ ID NO:3.  
XX  
XX  
SQ Sequence 516 AA;  
Query Match 99.9%; Score 2912; DB 21; Length 516;  
Best Local Similarity 99.8%; Pred. No. 1.3e-162;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MGVLVGLGALAGLGGFPAPAPGPGSQCVGHDGCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MGVLVGLGALAGLGGFPAPAPGPGSQCVGHDGCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGPGFQVMTGDNNTSYS 120  
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGPGFQVMTGDNNTSYS 120  
QY 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180  
DB 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180  
QY 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGCGCHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGCHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360  
DB 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360  
QY 361 ECVEPVDPFCFRANCEYQCPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
DB 361 ECVEPVDPFCFRANCEYQCPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
QY 421 TQASCEPEGYILDDGICHTIDCEGNGFCGSGVCHNLPGTFECICGPDLSALRHIGTDC 480  
DB 421 TQASCEPEGYILDDGICHTIDCEGNGFCGSGVCHNLPGTFECICGPDLSALRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPTPGSTLTTPAVGLVHSG 516  
DB 481 DSGKVDGDSGSGEPPTPGSTLTTPAVGLVHSG 516  
RESULT 7  
AAR11534  
ID AAR11534 standard; Protein; 575 AA.  
XX  
AC AAR11534;  
XX  
XX AAR11534;  
DT 17-JUN-1991 (first entry)  
XX  
DE Human thrombomodulin type II polypeptide.

XX Thrombomodulin; chondroitin; protein C; thrombin; thrombolysis;  
KW chondroitin sulphate; anticoagulant; myocardial infarction.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Peptide /label= signal peptide  
FT Peptide 19..365  
FT Peptide /label= opt. N-terminal sequence  
FT Peptide 366..480  
FT Peptide /label= human thrombomodulin active fragment  
FT Peptide 481..516  
FT Peptide /label= opt. C-terminal sequence  
XX  
XX WO9104276-A.  
XX  
XX 04-APR-1991.  
XX  
XX 25-SEP-1990; 90WO-JP01234.  
XX  
XX 25-SEP-1989; 89JP-0246270.  
XX  
XX (ASAH ) ASahi KASEI KOGYO.  
XX  
XX Yamamoto S, Gomi K, Ogawa K;  
XX  
XX WPI; 1991-117478/16.  
XX  
XX New human thrombomodulin polypeptide contg. chondroitin - has  
PT anticoagulant platelet agglutination and thrombolytic activity  
XX  
XX Disclosure; fig 1; 80pp; Japanese.  
XX  
XX This human thrombomodulin polypeptide comprises a 115 residue active  
CC fragment with opt. flanking sequences (N-terminal- 347 amino acids;  
CC C-terminal- 36 amino acids). Additionally it has an N-terminal signal  
CC sequence and an attached sugar chain, contg. chondroitin or chondroitin  
CC sulphate. It promotes the activation of protein C by thrombin and has  
CC good anticoagulant, platelet aggregation inhibition and thrombolytic  
CC activities. It can therefore be used for treating cardiovascular  
CC diseases, eg arteriosclerosis or myocardial infarction.  
XX  
XX  
SQ Sequence 575 AA;  
Query Match 99.9%; Score 2912; DB 12; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.5e-162;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MGVLVGLGALAGLGGFPAPAPGPGSQCVGHDGCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MGVLVGLGALAGLGGFPAPAPGPGSQCVGHDGCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGPGFQVMTGDNNTSYS 120  
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGPGFQVMTGDNNTSYS 120  
QY 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180  
DB 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180  
QY 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGCGCHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGCHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360  
DB 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360  
QY 361 ECVEPVDPFCFRANCEYQCPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420

```

|||||
Db 361 ECVEPVPDFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
Qy 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALARIHIGTDC 480
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
Qy 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516
Db 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516

RESULT 8
AAR41806
ID AAR41806 standard; peptide; 575 AA.
AC AAR41806;
XX
DT 30-MAR-1994 (first entry)
XX
Thrombomodulin.
KW Transformation; fungus; blood coagulation; prevention; platelet;
KW aggregation; thrombolytic activity; thrombolysis; human thrombomodulin;
KW site-directed mutagenesis; promotion; protein C; activation; thrombin.
XX
OS Homo sapiens.
XX
PN JP05213998-A.
XX
PD 24-AUG-1993.
XX
PF 02-AUG-1991; 91JP-0282369.
XX
PR 03-AUG-1990; 90JP-0204978.
PR 30-JUL-1991; 91JP-0189984.
XX
PA (ASAH ) ASAH CHEM IND CO LTD.
XX
WPI; 1993-299652/38.
XX
Novel polypeptide obtd. by culturing transformed fungus - having
PT blood coagulation preventing, platelet aggregation preventing and
PT thrombolytic activities
XX
Disclosure; Fig 55; 65pp; Japanese.
XX
Novel polypeptides, obtd. by culturing transformed fungus, have
CC blood coagulation preventing, platelet aggregation preventing
CC and thrombolytic activities.
CC In an example, plasmid M13mpl19MD3 (constructed from pSV2TMJ2
CC (ATCC 67238) contg. a human thrombomodulin sequence) was subjected
CC to site directed mutagenesis to prepare pSV2TMD7. Plasmid pSV2TMD7
CC was transfected to COS-1 cells. The activity of promoting protein C
CC activation by thrombin of the peptide produced by the transformed
CC COS-1 cell was measured. The amt. of the peptide was determined.
XX
SQ Sequence 575 AA;
Query Match 99.9%; Score 2912; DB 14; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.5e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MLGVILGALALAGLGPAPAEPPQGSQCQVHDCFCALYGPATFLNASQICDGLRGHLM 60
Db 1 MLGVILGALALAGLGPAPAEPPQGSQCQVHDCFCALYGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDKPLRGFLRGFWVGTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDKPLRGFLRGFWVGTGDNNTSYS 120
Qy 121 RWARDLNGAPLCGFLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAY 180

```

```

Db 121 RWARDLNGAPLCGFLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAY 180
Qy 181 EFGAAAAVSITYCTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 EFGAAAAVSITYCTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Qy 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAACQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAACQADGRSCTASATQSCNDLCEHFCVNP 300
Qy 301 DOPGSYSCMCEGYELAADQHRCEVDVDCILEPSPQRCQVNTQGGFECHCPYNDLVDG 360
Db 301 DOPGSYSCMCEGYELAADQHRCEVDVDCILEPSPQRCQVNTQGGFECHCPYNDLVDG 360
Qy 361 ECVEPVPDFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
Db 361 ECVEPVPDFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
Qy 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALARIHIGTDC 480
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
Qy 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516
Db 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516

RESULT 9
AAR43031
ID AAR43031 standard; protein; 575 AA.
XX
AC AAR43031;
XX
DT 16-MAY-1994 (first entry)
XX
Human thrombomodulin.
XX
Anticoagulant; platelet aggregation inhibitor;
KW protein C; activation; thrombin; thrombomodulin;
KW coagulation disorder; thrombosis; myocardial infarction;
KW embolism; telangiectasis; arteriosclerosis obliterans;
KW disseminated intravascular coagulation; DIC; angina pectoris;
KW gestosis; transient ischaemic attack.
XX
OS Homo sapiens.
XX
PN WO9322447-A.
XX
PD 11-NOV-1993.
XX
PF 30-APR-1993; 93WO-JP00578.
XX
PR 01-MAY-1992; 92JP-0112903.
XX
PA (ASAH ) ASAH CHEM IND CO LTD.
PA (ASAH ) ASAH KASEI KOGYO KK.
XX
Kondo S, Toma K, Zushi M;
XX
WPI; 1993-368806/46.
XX
Peptide with anticoagulant and platelet aggregation inhibitor
PT activity - which promotes protein C activation by thrombin and is
PT useful in treating coagulation disorders e.g. thrombosis
XX
Disclosure; Fig 1; 84pp; Japanese.
XX
New peptides (see AAR50069) are inhibitors of the blood coagulation
CC and platelet aggregation activities of thrombin and promote the
CC protein-C activation effect of thrombin. They can be produced
CC efficiently in pure form by culture of appropriate transformants,
CC and are useful in treatment of circulatory disorders such as
CC myocardial infarction, thrombosis, embolism, telangiectasis,

```

CC arteriosclerosis obliterans, disseminated intravascular  
CC coagulation, angina pectoris, gestosis and transient ischaemic  
XX attack.  
XX  
SQ Sequence 575 AA;

Query Match 99.9%; Score 2912; DB 14; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.5e-162;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVVLGALALAGLGFPAEPQPGSGQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MLGVVLGALALAGLGFPAEPQPGSGQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGRLGFWVTGDNNTSYS 120  
DB 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGRLGFWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180  
QY 181 EPGAAAAVSYTGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
DB 181 EPGAAAAVSYTGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300  
DB 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300  
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCQRVCVNTGGFECHCYPNYDLVDG 360  
DB 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCQRVCVNTGGFECHCYPNYDLVDG 360  
QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQTACPADCDPN 420  
DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDLSALARIHGTDC 480  
DB 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHGTDC 480  
QY 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPAVGLVHSG 516  
DB 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPAVGLVHSG 516

RESULT 10  
AAE23026  
ID AAE23026 standard; Protein: 575 AA.  
XX  
AC AAE23026;  
XX  
DT 21-AUG-2002 (first entry)  
XX  
DE Human thrombomodulin protein.  
XX  
KW Human; thrombomodulin; analogue; neurologic trauma; vulnery;  
KW spinal cord injury; neurological damage; therapy; tranquilliser;  
KW neuroprotective.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT /label= Signal\_peptide  
FT Protein 19..575  
FT /note= "Human mature thrombomodulin"  
FT Domain 1..226  
FT /note= "N-terminal domain"  
FT Domain 227..462  
FT /note= "6 EGF-like domain"  
FT Modified-site 463..497

FT Domain /note= "O-linked glycosylation site"  
FT 498..521  
FT /note= "Transmembrane domain"  
FT 522..557  
FT Domain /note= "Cytoplasmic domain"  
XX  
PN WO200217953-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 31-AUG-2001; 2001WO-US41930.  
XX  
PR 31-AUG-2000; 2000US-229714P.  
PR 23-AUG-2001; 2001US-093840S.  
XX  
PA (SCHD ) SCHERING AG.  
XX  
PI Festoff BW, Morser MJ;  
XX  
DR WPI; 2002-415551/44.  
DR N-PSDB; AAD36901.  
XX  
PT Treating neurological damage resulting from spinal cord injury in  
PT humans by administering a soluble, recombinant thrombomodulin analog  
PT which is resistant to oxidation -  
XX  
PS Claim 1; Fig 1; 3lpp; English.  
XX  
CC The invention relates to a method of using thrombomodulin analogues  
CC in the treatment of the neurologic trauma associated with spinal  
CC cord injury in mammals. The method involves administering a oxidation  
CC resistant, soluble, recombinant thrombomodulin analogue, where the  
CC methionine at position 388 has been replaced with a leucine, and the  
CC analogue is numbered in accordance with native thrombomodulin. The  
CC method and thrombomodulin analogue are useful for treating  
CC neurological damage resulting from spinal cord injury in human.  
CC The present sequence is human thrombomodulin protein.  
XX  
SQ Sequence 575 AA;

Query Match 99.9%; Score 2912; DB 23; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.5e-162;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVVLGALALAGLGFPAEPQPGSGQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MLGVVLGALALAGLGFPAEPQPGSGQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGRLGFWVTGDNNTSYS 120  
DB 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGRLGFWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180  
QY 181 EPGAAAAVSYTGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
DB 181 EPGAAAAVSYTGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300  
DB 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300  
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCQRVCVNTGGFECHCYPNYDLVDG 360  
DB 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCQRVCVNTGGFECHCYPNYDLVDG 360  
QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQTACPADCDPN 420  
DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDLSALARIHGTDC 480

```

|||||
Db 421 TQASCEPEGILDDGFICTDIDECENGFCGCHNLPGTFECICGPDLSALVRHIGTDC 480
Qy 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 11
AAE17521
ID AAE17521 standard; Protein; 575 AA.
AC AAE17521;
XX
XX 22-APR-2002 (first entry)
XX
XX Human full-length native thrombomodulin (TM).
XX
XX Human; thrombomodulin; TM analogue; protein C; thrombotic disease;
XX antithrombotic therapy; thrombin-activatable fibrinolysis inhibitor;
XX TAFI; myocardial infarction; disseminated intravascular coagulation;
XX DIC; deep vein thrombosis; septic shock; pulmonary embolism; angina;
XX acute respiratory distress syndrome; cancer; toxemia; septicemia;
XX circulatory disorder; coronary embolism; pulmonary embolism; stroke;
XX systemic coagulation disorder; immunosuppressive; cerebroprotective;
XX pregnancy; anticoagulant; thrombolytic; venous occlusive condition;
XX antibacterial; cardiant.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Peptide 1..18
XX /label= Signal_peptide
XX Protein 19..575
XX /note= "Human mature thrombomodulin"
XX Domain 19..244
XX /label= N-terminal_domain
XX Domain 245..480
XX /note= "6 EGF-like domains"
XX Modified-site 481..515
XX /note= "O-linked glycosylation site"
XX Domain 516..539
XX /label= Transmembrane_domain
XX Domain 540..575
XX /label= Cytoplasmic_domain
XX
XX WO200198352-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US19590.
XX
XX 21-JUN-2000; 2000US-213678P.
XX 12-JUN-2001; 2001US-0880484.
XX
XX (SCHD ) SCHERING AG.
XX
XX Light D, Morser MJ, Nagashima M;
XX
XX WPI; 2002-154622/20.
XX N-PSDB; AAD28182.
XX
XX Thrombomodulin analog with greater ability to potentiate
XX thrombin-mediated activation of protein C and lesser ability to
XX potentiate thrombin-mediated activation of thrombin-activatable
XX fibrinolysis inhibitor for treating thrombotic disease .
XX
XX Claim 1; Fig 4; 53pp; English.
XX
XX The invention relates to the design, production and use of thrombomodulin
XX (TM) analogues that have the ability to enhance the thrombin-mediated
XX activation of protein C which have a significantly reduced ability to
XX promote activation of thrombin-activatable fibrinolysis inhibitor (TAFI).

```

```

CC TM analogue is useful for treating thrombotic disease in human, by
CC administering it or a polynucleotide encoding it to the patient and
CC expressing the analogue in vivo. Composition comprising TM analogue is
CC useful for treating a thrombotic disease or condition in mammals. TM
CC analogue is useful in antithrombotic therapy. TM analogue is also useful
CC for treating diseases in which thrombus formation plays a significant
CC etiological role including myocardial infarction, disseminated
CC intravascular coagulation (DIC), deep vein thrombosis, septic shock,
CC pulmonary embolism, acute respiratory distress syndrome, unstable angina
CC and other arterial or venous occlusive conditions. TM analogue is also
CC useful for treating various circulatory disorders including coronary or
CC pulmonary embolism, stroke and systemic coagulation disorders including
CC DIC which is often associated with septicemia, certain cancers and
CC toxemia of pregnancy. The present sequence is human full-length native
CC thrombomodulin.
XX
XX Sequence 575 AA;
XX
XX Query Match 99.9%; Score 2912; DB 23; Length 575;
XX Best Local Similarity 99.8%; Pred. No. 1.5e-162;
XX Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MLGVLVLGALALAGLGFAPAPBPQGGSCQVEHDCFALYPPGPFALNASQICDGLRGHLM 60
XX DB 1 MLGVLVLGALALAGLGFAPAPBPQGGSCQVEHDCFALYPPGPFALNASQICDGLRGHLM 60
XX
XX QY 61 TVRSSVAADVISILLNGDGGVGRRLWIGLQLPGGGDPKRLGPGFOWVTGDNNTSYS 120
XX DB 61 TVRSSVAADVISILLNGDGGVGRRLWIGLQLPGGGDPKRLGPGFOWVTGDNNTSYS 120
XX
XX QY 121 RWARLDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQCCVEKADGFLCEHFFPATCRPLAV 180
XX DB 121 RWARLDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQCCVEKADGFLCEHFFPATCRPLAV 180
XX
XX QY 181 EPGAAAASVITYGTPTFAARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240
XX DB 181 EPGAAAASVITYGTPTFAARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240
XX
XX QY 241 GAWDCSVENGCGCEHACNAIPGAPRCOCAGALOADGRSCTASATOSCNDCLEHFCVNP 300
XX DB 241 GAWDCSVENGCGCEHACNAIPGAPRCOCAGALOADGRSCTASATOSCNDCLEHFCVNP 300
XX
XX QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQPCRCVNTQGGFECHECHYPNYDLVDG 360
XX DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQPCRCVNTQGGFECHECHYPNYDLVDG 360
XX
XX QY 361 ECVEPVDPFCFRANCEYQCOPLNQTSYLCVCAEGFADIPHEPHRCQMFQACPDADCPN 420
XX DB 361 ECVEPVDPFCFRANCEYQCOPLNQTSYLCVCAEGFADIPHEPHRCQMFQACPDADCPN 420
XX
XX QY 421 TQASCECEGYILDDGFICTDIDECENGFCGCHNLPGTFECICGPDLSALVRHIGTDC 480
XX DB 421 TQASCECEGYILDDGFICTDIDECENGFCGCHNLPGTFECICGPDLSALVRHIGTDC 480
XX
XX QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
XX DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
XX
XX RESULT 12
XX AAR10617
XX ID AAR10617 standard; Protein; 515 AA.
XX
XX AC AAR10617;
XX
XX DT 17-APR-1991 (first entry)
XX
XX DE Soluble thrombomodulin deriv.
XX
XX KW Thrombosis; anticoagulant; Protein C.
XX
XX OS Homo sapiens.
XX

```

FH Key Location/Qualifiers
FT Peptide 1..16
FT /label= signal peptide
FT /note= "for low mol. wt deriv.; optionally present"
FT Peptide 1..18
FT /label= signal peptide
FT /note= "for high mol wt. deriv; optionally present"
XX
XX EP412841-A.
XX
XX 13-FEB-1991.
XX
XX 10-AUG-1990; 90EP-0308826.
XX
XX 05-FEB-1990; 90US-0474870.
XX 11-AUG-1989; 89US-0393617.
XX
XX (ELIL ) ELI LILLY & CO.
XX
XX Bang NU, Grinnell BW, Hoskins JA, Moore RE, Parkinson JF;
XX
XX WPI: 1991-045960/07.
XX N-PSDB; AAQ10435.
XX
XX New soluble thrombomodulin derivatives - produced by recombinant
PT DNA techniques for use as anticoagulants and in treatment of
PT thrombosis.
XX
XX Claim 2; Page 45; 81pp; English.
XX
XX The sequence is that of a sol. thrombomodulin deriv. comprising the
CC signal peptide (optional), N-terminal, epidermal growth factor
CC homology region and the Ser/Thr-rich region, but lacking the trans-
CC membrane and cytoplasmic domains. The sequence was deduced DNA
CC obt'd. from clone GHM3A isolated from a human chromosome 20 library,
CC subcloned into pUC19 to give pGHTM3A. A Pvu MI fragment of this
CC vector was ligated with a linker to construct pUC18TM, which was
CC treated with BsmI to delete about 500 bp from the 3' end of the
CC gene. ( The deletion occurs at the Ser/Thr-rich/transmembrane domain
CC junction, effectively cleaving the transmembrane and cytoplasmic
CC domains.) The plasmid was recircularised to give pUC18TMd for use
CC in prodn of the recombinant sol. deriv. The protein reacts with
CC thrombin to activate the Protein C anticoagulant pathway, inhibit
CC thrombin, and can be used to treat or prevent thrombotic disorders.
XX
XX Sequence 515 AA;
XX
XX Query Match 99.8%; Score 2910; DB 12; Length 515;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-162;
XX Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MLGVLVLGALALAGLGFPAPEPQPGSGQCVHDCFALYGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFPAPEPQPGSGQCVHDCFALYGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGFGQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGFGQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGPLCAVSAEAATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGPLCAVSAEAATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAV 180
QY 181 EPGAAAAVSTYGTPTFAARGADQALPVGSSAAVAPLGLQLMCTAPPAGVOGHWAREAP 240
DB 181 EPGAAAAVSTYGTPTFAARGADQALPVGSSAAVAPLGLQLMCTAPPAGVOGHWAREAP 240
QY 241 GAWDCSVENGCCERHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCCERHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLYDG 360

Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLYDG 360
QY 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFCNOTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFCNOTACPADCDPN 420
QY 421 TQASCECPGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPPSALARHIGTDC 480
Db 421 TQASCECPGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPPSALARHIGTDC 480
QY 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHS 515
Db 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHS 515
RESULT 13
AAP82070
ID AAP82070 standard; protein; 575 AA.
XX
XX AAP82070;
XX
XX 19-OCT-1990 (first entry)
XX Human thrombomodulin encoded by plasmid p2.1.
XX
XX thrombomodulin activity; Protein C; anticoagulant;
KW epidermal Growth Factor (EGF) domains.
XX
XX synthetic.
XX OS
XX WO8809811-A.
XX
XX 15-DEC-1988.
XX
XX 09-JUN-1988; 88WO-DK00089.
XX
XX 12-JUN-1987; 87DK-0002990.
XX
XX (NOVO ) NOVO INDUSTPRI A/S.
XX
XX Nexo BA, Esper B;
XX
XX WPI: 1988-368626/51.
XX N-PSDB; AAP82026.
XX
XX Recombinant protein having thrombomodulin activity -
PT used in the therapeutic control of coagulation and the treatment
PT and prevent of thrombotic episodes
PT
XX Disclosure; ; p; English.
XX
XX Protein is encoded by plasmid p2.1 derived from human cell
CC line A549 known to express about 10000 molecules of thrombomodulin
CC per cell. Plasmid p2.1 showed a strong hybridisation signal with a
CC 60-mer bovine thrombomodulin probe. Thrombomodulins are used to
CC potentiate a patient's anticoagulant capacity.
CC See also AAN82027 and AAN82037.
XX
XX Sequence 575 AA;
XX
XX Query Match 99.8%; Score 2910; DB 9; Length 575;
XX Best Local Similarity 99.8%; Pred. No. 1.9e-162;
XX Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLGVLVLGALALAGLGFPAPEPQPGSGQCVHDCFALYGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFPAPEPQPGSGQCVHDCFALYGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGFGQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGFGQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGPLCAVSAEAATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAV 180



```
|||||
Db 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADAFCEHFFPATCRPLAV 180
QY 181 EPGAAAASVTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAASVTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
QY 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPPNP 300
Db 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPPNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVPEVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN 420
Db 361 ECVPEVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN 420
QY 421 TQASCECEGYILDGFICTDIDECENGFCGSGVCHNLPCTFECICGPDSSALARHIGTDC 480
Db 421 TQASCECEGYILDGFICTDIDECENGFCGSGVCHNLPCTFECICGPDSSALARHIGTDC 480
QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 14
AAR22018
ID AAR22018 standard; Protein; 516 AA.
XX AC AAR22018;
XX DT 03-JUL-1992 (first entry)
XX DE Human thrombomodulin (1-516) with Asp367 substituted by Glu.
XX KW Mutant; thrombin binding site; blood clotting; Tm2 mutator.
XX OS Homo sapiens.
XX PN EP474273-A.
XX PD 11-MAR-1992.
XX PF 05-AUG-1991; 91EP-0202009.
XX PR 03-AUG-1990; 90JP-0204978.
XX PA (ASAH ) ASAH KASEI KOGYO.
XX Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;
XX WPT; 1992-081820/11.
XX PT New polypeptide inhibits blood coagulation and platelet
XX aggregation - promotes thrombin catalysed protein C activation
XX for treating myocardial infarction, thrombosis, embolism, etc.
XX PS Example 1; Page 18; 112pp; English.
XX CC Plasmid M13TMD1 (see AAR22016) encoding the first 516 N-terminal
XX amino acids of human thrombomodulin was used as a template for
XX site-directed mutagenesis using the "Tm2 mutator" to produce
XX plasmid M13TMD2. In the mutant thrombomodulin encoded by the
XX plasmid, the wild-type Asp residue at position 367 is substituted
XX by an Glu residue. The activity of this mutant, truncated
XX thrombomodulin (i.e. D123glu) in the activation of protein C was
XX compared to that of similarly truncated thrombomodulin with Asp
XX at position 367 (i.e. D123asp). The activity of D123glu was as high
XX as 1.2 times that of D123asp. See AAR22013-R22022 and AAR25074.
```

```
SQ Sequence 516 AA;
Query Match 99.7%; Score 2908; DB 13; Length 516;
Best Local Similarity 99.6%; Pred. No. 2.3e-162;
Matches 514; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLGVLVGALALAGLGFPAAPQPGGQCVEHDCFALYGPATFFLNASICDGLRHLM 60
Db 1 MLGVLVGALALAGLGFPAAPQPGGQCVEHDCFALYGPATFFLNASICDGLRHLM 60
QY 61 TVRSSVAADVLSILLNGDGGVRRRLWIGLQPPGCCDPRKLGFRGFQWVTGNNTSYS 120
Db 61 TVRSSVAADVLSILLNGDGGVRRRLWIGLQPPGCCDPRKLGFRGFQWVTGNNTSYS 120
QY 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
Db 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
QY 181 EPGAAAASVTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAASVTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
QY 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPPNP 300
Db 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPPNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVPEVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN 420
Db 361 ECVPEVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN 420
QY 421 TQASCECEGYILDGFICTDIDECENGFCGSGVCHNLPCTFECICGPDSSALARHIGTDC 480
Db 421 TQASCECEGYILDGFICTDIDECENGFCGSGVCHNLPCTFECICGPDSSALARHIGTDC 480
QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 15
AAR31572
ID AAR31572 standard; Protein; 575 AA.
XX AC AAR31572;
XX DT 27-MAY-1993 (first entry)
XX DE Human thrombomodulin.
XX KW Mutagenesis; site specific; nucleic acid constructs;
XX restriction site; introduction; removal.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Peptide 1..20
FT Domain /note= "signal peptide"
FT Domain 245..280
FT Domain /note= "EGF-1"
FT Domain 288..323
FT Domain /note= "EGF-2"
FT Domain 329..361
FT Domain /note= "EGF-3"
FT Domain 368..404
FT Domain /note= "EGF-4"
FT Domain 408..439
FT Domain /note= "EGF-5"
FT Domain 445..480
FT Domain /note= "EGF-6"
```



Search completed: December 16, 2002, 17:24:25  
Job time : 35.5 secs

```
FT Domain 481..515
FT Region /note="O-linked glycosylation domain"
FT 516..538
FT /note="stop transfer sequence"
FT 539..575
FT /note="cytoplasmic domain"
XX
PN WO9301282-A.
XX
XX 21-JAN-1993.
XX
XX 01-JUL-1992; 92WO-US05573.
XX
XX 01-JUL-1991; 91US-0724237.
XX
XX (BERL-) BERLEX LAB INC.
XX
XX Andrews WH, Morser MJ, Vilander LR;
XX
XX WPI; 1993-045488/05.
XX
XX Site-specific mutagenesis of nucleic acid constructs - using an
PT oligo:nucleotide which changes a nucleotide and introduces or
PT removes a restriction site
XX
XX Example; Fig 3; 87pp; English.
XX
XX The sequence is that of the native human thrombomodulin showing the
CC six EGF-like domains.
XX
XX Sequence 575 AA;

Query Match 99.7%; Score 2908; DB 14; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.5e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGVLVIGALALAGLGPAPAEPOGGSCQVEHDCPALYFGPATFLNASQICDGLRGHLM 60
Db 1 MGVLVIGALALAGLGPAPAEPOGGSCQVEHDCSALYFGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVADVTISLLNGDGGVRRRLWIGLQPPCGDKPLRGPLRGFWVTGDNNTSYS 120
Db 61 TVRSSVADVTISLLNGDGGVRRRLWIGLQPPCGDKPLRGPLRGFWVTGDNNTSYS 120
Qy 121 RWARLDLNGAPLCGLPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
Db 121 RWARLDLNGAPLCGLPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
Qy 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAGVGHWAREAP 240
Db 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAGVGHWAREAP 240
Qy 241 GAWDCSVENGGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Qy 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILESPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILESPSPQRCVNTQGGFECHCYPNYDLVDG 360
Qy 361 ECVPEVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVPEVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Qy 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSDALARRHIGTDC 480
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSDALARRHIGTDC 480
Qy 481 DSGKVDGSDSGSGEPSPPTFGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSGEPSPPTFGSTLTTPPAVGLVHSG 516
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:22:58 ; Search time 16.5 Seconds  
(without alignments)  
3006.386 Million cell updates/sec

Title: US-09-509-994-2  
Perfect score: 2916  
Sequence: 1 MLGVLVGLALAGLGFAP.....PSTPGSTLTPPAVGLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2912	99.9	575	1 THHUB	thrombomodulin pre
2	1903	65.3	577	2 A60501	thrombomodulin pre
3	1095.5	37.6	356	2 A25918	thrombomodulin - b
4	414	14.2	1574	2 T13954	MEGF6 protein - ra
5	414	14.2	1620	2 T12783	hypothetical prote
6	374	12.8	2907	2 A57278	fibrillin-2 precu
7	368	12.6	2871	2 A55367	fibrillin-1 - bovi
8	367.5	12.6	1184	2 A55184	fibulin-2 precu
9	366	12.6	2918	2 A54105	fibillin-2 precu
10	364	12.5	3002	2 A47221	fibillin-1 precu
11	362	12.4	2871	2 A55624	fibillin-1 precu
12	357	12.2	741	2 T46488	hypothetical prote
13	357	12.2	1221	2 A49457	fibulin-2 precu
14	312.5	11.4	1964	2 T09059	notch4 - mouse
15	331	11.4	589	2 T43210	fibulin-1D precu
16	330.5	11.3	1712	2 A38261	masking protein pr
17	324.5	11.1	689	2 T42760	fibulin 1, splice fo
18	324.5	11.1	712	2 T42990	fibulin 1, splice
19	322.5	11.1	1394	2 A35626	transforming growt
20	321	11.0	2321	2 S78549	notch3 protein - h
21	314.5	10.8	3507	2 T34513	hypothetical prote
22	313	10.7	798	2 T22793	hypothetical prote
23	312.5	10.7	1820	2 A55494	latent transformin
24	311	10.7	601	2 B36346	fibulin 1 precu
25	311	10.7	683	2 C36346	fibulin 1 precu
26	309	10.6	685	2 S78040	fibulin, splice fo
27	308	10.6	705	2 S34968	fibulin, splice fo
28	303.5	10.4	1251	2 A57293	latent transformin
29	297	10.2	2531	2 A46019	Notch-1 protein -

RESULT 1

THHUB

thrombomodulin precursor [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1988 #sequence\_revision 12-May-1995 #text\_change 15-Sep-2000  
C:Accession: A41442; A28307; A29680; A27073; JX0264; S3954

R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruy

J. Biochem. 103, 281-285, 1988

A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed ac

A:Reference number: A41442; MUID:88227901; PMID:2836377

A:Accession: A41442

A:Molecule type: DNA

A:Residues: 1-575 <SHI>

A:Cross-references: DDBJ:D00210; NID:g220126; PIDN:BAA00149.1; PID:g220127

R:Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.

Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987

A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of the

A:Reference number: A28307; MUID:87317665; PMID:2819876

A:Accession: A28307

A:Molecule type: DNA; mRNA

A:Residues: 1-472, 'A', 474-575 <JAC>

A:Cross-references: GB:J02973; NID:g339658; PIDN:AAA61175.1; PID:g339659

R:Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioaka, J.; Maruyama, I.; Zushi, M.; Kaw

EMBO J. 6, 1891-1897, 1987

A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on end

A:Reference number: A29680; MUID:88004395; PMID:2820710

A:Accession: A29680

A:Molecule type: mRNA

A:Residues: 1-575 <SUZ>

A:Cross-references: GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:g736251

A:Experimental source: lung endothelium

A:Note: Part of this sequence, including the amino end of the mature protein, were de

R:Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.

Biochemistry 26, 4350-4357, 1987

A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of

A:Reference number: A27073; MUID:88024950; PMID:2822087

A:Accession: A27073

A:Molecule type: mRNA

A:Residues: 1-472, 'A', 474-575 <WEN>

A:Cross-references: GB:M16552; NID:g339656; PIDN:AAB59508.1; PID:g339657

A:Experimental source: placenta

A:Note: Parts of this sequence were determined by protein sequencing

R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.

J. Biochem. 113, 433-440, 1993

A:Title: Urinary thrombomodulin, its isolation and characterization.

A:Reference number: JX0264; MUID:93293792; PMID:8390446

A:Accession: JX0264

A:Molecule type: protein; mRNA

A:Residues: 19-472, 'A', 474-486 <YAM>

A:Experimental source: urine

A:Note: the urinary form appears to be identical with that circulating in plasma

R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.

Biochem. J. 295, 131-140, 1993  
A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble serine.  
A:Reference number: S38954; MUID:94029900; PMID:8216207  
A:Accession: S38954  
A:Molecule type: protein  
A:Residues: 475-491, 'X', 493-494 <GER>  
A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chondroitin-6-sulfate.  
R:Maininger, D.P.; Komives, E.A.  
submitted to the Brookhaven Protein Data Bank, September 1995  
A:Reference number: A67369; PDB:1ZAO  
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
R:Tulinsky, A.; Mathews, I.I.  
submitted to the Brookhaven Protein Data Bank, August 1994  
A:Reference number: A52804; PDB:1HLL  
A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442  
R:Hrabal, R.; Komives, E.A.; Ni, F.  
submitted to the Brookhaven Protein Data Bank, November 1995  
A:Reference number: A63583; PDB:1FGD  
A:Contents: annotation; conformation by (1)H-NMR, residues 427-444  
R:Hrabal, R.; Komives, E.A.; Ni, F.  
Protein Sci. 5, 195-203, 1996  
A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, thrombospondin-1.  
A:Reference number: A58595; MUID:96276211; PMID:8745396  
A:Contents: annotation; conformation by (1)H-NMR  
C:Genetics:  
A:Gene: GDB:THBD  
A:Cross-references: GDB:119613; OMIM:188040  
A:Map position: 20p11.2-20p11.2  
A:Introns: \*status absent  
C:Complex: homodimer, urinary form  
C:Function:  
A:Description: Inhibits thrombin activation of fibrinogen; cofactor for thrombin activation of plasminogen.  
A:Pathway: blood coagulation/moderation  
A:Note: the membrane-bound form is located on the endothelium luminal surface of arteries.  
A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis.  
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C:Keywords: anticoagulant; beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coagulation; protein  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>  
F:19-513/Domain: extracellular #status predicted <EXT>  
F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>  
F:24-167/Domain: C-type lectin homology <LCH>  
F:177-199/Region: PEST sequence  
F:201-233/Region: PEST sequence  
F:245-280/Domain: EGF homology <EG1>  
F:288-323/Domain: EGF homology <EG2>  
F:329-362/Domain: EGF homology <EG3>  
F:369-404/Domain: EGF homology <EG4>  
F:408-439/Domain: EGF homology <EG5>  
F:445-480/Domain: EGF homology <EG6>  
F:485-513/Region: PEST sequence  
F:517-539/Domain: transmembrane #status predicted <TMN>  
F:540-575/Domain: intracellular #status predicted <INT>  
F:47,115,116,382,409/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:174,225,411,504/Binding site: carbohydrate (Thr) (covalent) #status predicted  
F:245-256,252-265,267-280,288-296,292-308,310-323,329-340,336-349,351-362,369-378,374-388/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F:334,498/Binding site: erythro-beta-hydroxyasparagine (Asn) #status experimental  
F:342/Modified site: chondroitin sulfate (Ser) (covalent) (partial) #status experimental  
F:490,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimental

Query Match 99.9%; Score 2912; DB 1; Length 575;  
Best Local Similarity 99.8%; Pred. No. 2.8e-166;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLGVLVIGALAGLGPAPAEQPGSGQVHDCFCALYGPATFNLASQICDGLRGHLM 60  
Db 1 MLGVLVIGALAGLGPAPAEQPGSGQVHDCFCALYGPATFNLASQICDGLRGHLM 60

Qy 61 TVRSSVAADVISLLNDDGGVRRRLWIGLQLPPGGDKPLRGFLRGFWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNDDGGVRRRLWIGLQLPPGGDKPLRGFLRGFWVTGDNNTSYS 120

Qy 121 RWRDLNGAPLCPLCVAVSAAEATVPSPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180  
Db 121 RWRDLNGAPLCPLCVAVSAAEATVPSPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180  
Qy 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVAVQGHAREAP 240  
Db 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVAVQGHAREAP 240  
Qy 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATSCNDLCHEFCVPNP 300  
Db 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATSCNDLCHEFCVPNP 300  
Qy 301 DQGSYSQMCETGYRLAADQHRCEVDVDCILEFSPQRCVNTQGGFECYCNIDLVGG 360  
Db 301 DQGSYSQMCETGYRLAADQHRCEVDVDCILEFSPQRCVNTQGGFECYCNIDLVGG 360  
Qy 361 ECVEPVPDFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQTACPADCPN 420  
Db 361 ECVEPVPDFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQTACPADCPN 420  
Qy 421 TQASCECEGYLDDGFICTDIDECENGCGFCGVCHNLPCTFECICGSDSALARHIGTDC 480  
Db 421 TQASCECEGYLDDGFICTDIDECENGCGFCGVCHNLPCTFECICGSDSALARHIGTDC 480  
Qy 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 2  
A60501  
thrombomodulin precursor - mouse  
N:Alternate names: fetomodulin  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Jul-1999  
C:Accession: S08488; A32001; A60501  
R:Dittman, W.A.; Majerus, P.W.  
Nucleic Acids Res. 17, 802, 1989  
A:Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted  
A:Reference number: S08488; MUID:89128454; PMID:2536925  
A:Accession: S08488  
A:Molecule type: mRNA  
A:Residues: 1-577 <DIT>  
A:Cross-references: EMBL:X14432; NID:g54781; PID:CAA32597.1; PID:g54782  
R:Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.  
J. Biol. Chem. 263, 15815-15822, 1988  
A:Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetate  
A:Reference number: A32001; MUID:89008498; PMID:2844823  
A:Accession: A32001  
A:Molecule type: mRNA  
A:Residues: 97-577 <DI2>  
A:Cross-references: CB:J04060  
R:Imada, S.; Yamauchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.  
Dev. Biol. 140, 113-122, 1990  
A:Title: Identification of fetomodulin, a surface marker protein of fetal development  
A:Reference number: A60501; MUID:90292331; PMID:2162790  
A:Accession: A60501  
A:Molecule type: protein  
C:Residues: 19-22; 330-343; 479-489; 545-555; 562-575 <IMA>  
C:Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofactor  
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C:Keywords: anticoagulant; membrane protein; phosphoprotein; receptor  
F:24-165/Domain: C-type lectin homology <LCH>  
F:244-279/Domain: EGF homology <EG1>  
F:287-322/Domain: EGF homology <EG2>  
F:328-361/Domain: EGF homology <EG3>  
F:368-403/Domain: EGF homology <EG4>  
F:407-438/Domain: EGF homology <EG5>  
F:444-479/Domain: EGF homology <EG6>

Query Match 65.3%; Score 1903; DB 2; Length 577;  
Best Local Similarity 65.6%; Pred. No. 3.1e-106;

Matches 342; Conservative 46; Mismatches 125; Indels 8; Gaps 4;

Qy	1	MLGVVLGALALACIGFPAPAEPOPGSGSQCVHEHDCFALYPGPATFLNASQICDCLGRHLM	60
Db	1	MLGIFFLVGLAPASGLSALAKLQPTGSGQVEHECFALFQGPATFLDASQACORLQGHLM	60
Qy	61	TVRSSVAADVISILLNGDG -GVGRRRLRWIGLQLPGGCDPKRLGFGQVWTGDNNTSY	119
Db	61	TVRSSVAADVISILLSSQSMDLGP ---WIGLQLPQGCDDPVHLGFLRGFGQVWTGDNNTSY	117
Qy	120	SRWARLDNGALPCGGLPCLVAVSAEAATVPSEPTWEEQOCEKADGFLCEHFHPTACRPLA	179
Db	118	SRWARPNQDAPLPCGGLPCLVTVSTATEAAPGEPAWEKPCETETGFLCEFFYFTASCRPLT	177
Qy	180	VEP -GAAAAAVSYTYGPPFAARGADFOALPVGSSAAVAPLQJLMCTAPPAQVQGHWARE	238
Db	178	VNTRDPEAAHISSTYNTPFGVSGADFOTLPVGSSAAVEPLGGLVLCRAPPGTSEGHAWAE	237
Qy	239	APGAWDCSVENGGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVP	298
Db	238	ATGAWNCVSENGGCEYLCNRSTNEPRCLCPRDMQLQADGRSCARPVVGSCNELCEHFCVS	297
Qy	299	NPDPQPGSYSCMCEYGYRLAADHCECDVDDCILEPSPCQPCVNTQGGFECYCPNYDLV	358
Db	298	NAEYPPGYSYSCMCEYGYQLAADHGHCEDDVDDCKQGNPCPQLCVNTKGFEFCYDGYELV	357
Qy	359	DGECVEPVDPCFRANCEYQCOPLNOTSYLCVCABGFAPIPHEPHRCQFCNQATCAPDCD	418
Db	358	DGECVELLDPCFGNCEFCQCPVSPDYRCICAPGFAPKDPBPHKCEMFCNETSCPADC	417
Qy	419	PNTQASCEPBGYITLDGFICTIDICEENGFCSGVCHNLPGTFPECICGPPSALARHIGT	478
Db	418	PNSPTVCECPGFIILDEGSVCTDIDBCSQGCEFTSECRNFGPSYECICGPDIALAGQISK	477
Qy	479	DCDSKGV ---DGGDSGSGGEPSPPTGSLTTPPAVGLVHSG	516
Db	478	DCDPTPVREDPKKEEGSGEPSPVSPGSGTGPSPARPVHSG	518

### RESULT 3

A25918  
thrombomodulin - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 16-Jul-1999  
C:Accession: A25918  
R:Jackman, R.W.; Beeler, D.L.; VanDeWater, L.; Rosenberg, R.D.  
Proc. Natl. Acad. Sci. U.S.A. 83, 8834-8838, 1986  
A:Title: Characterization of a thrombomodulin cDNA reveals structural similarity to the  
A:Reference number: A25918; MUID:87067408; PMID:3024152  
A:Accession: A25918  
A:Molecule type: mRNA  
A:Residues: 1-356 <JAC>

A; Residues: 1-356 <JB>  
A; Cross-references: GB<M14657> NID:gl63762; PIDN:AA30785.1; PID:gl63763  
C; Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C; Keywords: transmembrane protein  
E; 21-56/Domain: EGF homology <EG1>  
F; 64-97/Domain: EGF homology <EG2>  
F; 103-136/Domain: EGF homology <EG3>  
F; 143-178/Domain: EGF homology <EG4>  
F; 182-213/Domain: EGF homology <EG5>  
F; 219-253/Domain: EGF homology <EG6>

Query Match	37.6%	Score 1095.5;	DB 2;	Length 356;
Best Local Similarity	63.4%;	Pred. No. 2.3e-58;		
Matches 189; Conservative	31;	Mismatches 63;	Indels 15;	Gaps 6

229	GAVOGHWAREAPGAWDCSVENGCGEHCACNAIPGAPRCQCPAGAALQADGRCSCTASATQSC	288
QY		
5	GETEGRSKEAPGAWACGVERGCGHECKGSAGASNCLCPADAALQADGRCSGLPAEHPC	64
DB		
289	NDLCHFCVCPNPDQPGSYSCMCETGYRLAADQHRCEDEVDDCILEPSPCORCVNTQGGFE	348
QY		
65	HOICPEFFCT - HUHGIGNTVTCICEAGYLAADQHRCEDEVDDCAOLPSPCORCVNTQGGFQ	122
DB		

db 65 HOLCEHEC--HLHGLGNYTCICEAGYOLAAADQHRCEDVDDCAQLPSPCPQRCVNTGGFQ 122

QY	349	CHCYPNYDLVDGCEVPVDFCFRANCEYQCQPLNQTSYLCVCAGGPAPTPHPRHCQMFC	408
		:       :       :       :       :       :       :       :       :	
Db	123	CHCDTGYELVDGECVDPVDFCNNCCEYQCQPVSEHKICIAEGFAPVPGAPHKQMFC	182
QY	409	NQTACPADCDPNTQAACEPEYLDDGFCTDIDECENGGFGCSGVCHNLPGTFEFCIGCP	468
		:       :       :       :       :       :       :       :       :	
Db	183	NQTSCPADCDPHPTTCRCPEYIIDEGSTCTDINECDT-NICPGQCHNLPGTYEFCIGCP	241
QY	469	DSALARHIGTDCDSGKV-----DGGDSGSCEPPSPPTPGSLTLP-PA-VCLVHS	516
Db	242	DSALSGOIGTDCDPTOVNEERGTPEDYG--GSGEPVSPPTPGATARPSPAPAGPLHSG	297

## RESULT 4

Tl13954  
MEGF6 protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: Tl13954  
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A:title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A:Reference number: Z14126; MUID:98360089; PMID:9693030  
A:Accession: Tl13954  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-1574 <NA>  
A:Cross-references: EMBL:AB011532; NID:G3449293; PIDN:BAA32462.1; PID:G3449294  
A:Experimental source: strain Sprague-Dawley; brain  
C:Genetics:  
A:Gene: MEGF6

Query Match 14.2%; Score 414; DB 2; Length 1574;  
Best Local Similarity 34.1%; Pred. No. 2.7e-17;  
Matches 104; Conservative 28; Mismatches 101; Indels. 72; Gaps 17;

[illegible]

## RESULT 5

T27283  
hypothetical protein Y6G10A.f - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T27283  
R:Ainscough, R.  
submitted to the EMBL Data Library, September 1999  
A:Reference number: Z20336  
A:Accession: T27283  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-1620 <WTL>  
A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CESP:Y64G10A.f  
A:Experimental source: clone Y64G10A  
C:Genetics:  
A:Gene: CESP:Y64G10A.f  
A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 548/1; 559/1; 601/1; 625/1;  
Query Match 14.2%; Score 414; DB 2; Length 1620;  
Best Local Similarity 27.0%; Pred. No. 2.7e-17;  
Matches 128; Conservative 38; Mismatches 176; Indels 132; Gaps 20;  
QY 119 YSRWRLDLNGLPLGCLVAVSAE-----ATVSEPTWEEOQCE--VKADGFCEP 169  
DB 56 YLRFARFRSGCKCLLRVQANCSADLCHNGGTCVPSEHNDNEQVCEPCFTGACQY 115  
QY 170 HPPATCRPLAVPFGAAAVSTVGTTPFAARGADFOALVGVSSAAVAPLGLQLMCTAPPG 229  
DB 116 D-ANEC--MANNGGCEHECVN-TIGTYI-----CRCWPG 145  
QY 230 AVOGHWAEPAGWDCSYVNGGCEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSC- 288  
DB 146 FELSGDNGTCSDDIDECVNSGCDRCVNSPGFRCDPCSDLYLHADGRTG--GKVTSCS 203  
QY 289 --NDLCEHFCVPPNPQPGSYSCMCTGYRLAADOHRCEVDVDCILEPSPQRCVNTQGG 346  
DB 204 TDNGGCEHC--ENDSNGEFYRCRCVGFKLSENKRSQVDPVPCFDNKGCGQHCTNNHR 262  
QY 347 FCHCYPNVDL--VDCCEVEPVDPFRAN--CEYQCOPLNOTSYLVCVCAEGFAPIPHEPHRC 404  
DB 263 ACQCYPGFHLISYDRSCVDIDECANNGCEHFCENKGT--YRCKREGY--QLGRDGRTC 320  
QY 405 QMF-----CNQACPADC--DPTQASCEPCEGYL----- 433  
DB 321 EEMLGCGVNGGCGQHQDCYDQDGGHVCKCRNGYILANDQKLCHDNTSTVIHARAPLWD 380  
QY 434 -----DDGF----- 455  
DB 381 SYETVTCVPTDITLTKLCHLMDSHGVQCFDGDGYELIDSFQCDIINFCHENNGDSCQIC 440  
QY 456 HNLPGTFECICGPDSSALAHITG-----DCDS-----GKVDGDSGSGEP 495  
DB 441 VNLGVSCEQCKPFLMKRKTCEIDSECSNNGGCGQICSNQEGGYMCSCEP 494  
RESULT 6  
A57278  
fibrillin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 02-Aug-2002  
C:Accession: A57278  
R:Zhang, H.; Hu, W.; Ramirez, F.  
J. Cell Biol. 129, 1165-1176, 1995  
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extracellular matrix  
A:Reference number: A57278; MUID:95263670; PMID:7744963  
A:Accession: A57278  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2907 <ZHA>  
A:Cross-references: GB:L39790; NID:g762830; PIDN:AAA74908.1; PID:g762831  
C:Superfamily: fibrillin 1; EGF homology  
F:1239-1274/Domain: EGF homology <EGF>  
F:2488-2523/Domain: EGF homology <EGF>  
Query Match 12.8%; Score 374; DB 2; Length 2907;  
Best Local Similarity 32.9%; Pred. No. 1.1e-14;  
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;  
QY 244 DCSVNGGCEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNLCHEFCVNPDD-- 301  
DB 1238 ECMIMNGGCDTQCTNSGSEYECSCSEGYALMPDGRSCA-----DIDE--CENNPDIC 1287  
QY 302 -----QPGSYSCMCTGYRLAADOHRCEVDVDCILEPSPCP--QRCVNTQGGFECHCYP 353

DB 1288 DGGQCTNIPGEYRCLCYDGFMA5MDKMTCLDVNECDLNENICMFGECENTKGSFICHCOL 1347  
QY 354 NYDLVDGE--CVBPVDPG--FRANCEYQCOPLN-QTSYLCVCAEGFA-----PIP 398  
DB 1348 GYSVKGGTGTCTD-VDECEIGHNCDMHASCLNVPGSFKSCREGWVGNGIKCIDLDECA 1406  
QY 399 HEHRCQMFNCTACPADCDPNTQAS--CECPGYILLDGGFTCTDIDE-----CENG- 448  
DB 1407 NGTHQCSI-----NAQC--VNTPGSYRCACSEGF-TGDFGFTCSVDDECAENTLNCENQ 1457  
QY 449 -----GF-----CS-----GVCHNLPCTFECICGPDSSALA 473  
DB 1458 CLNVPAGYRCECEMGFTPASDSRSCQIDECSPQNCVFTGTCNNLPGMFHCICDDGYELD 1517  
QY 474 RHIG--TDICD 481  
DB 1518 RTGGNCTDID 1527  
RESULT 7  
A55567  
fibrillin 1 - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Aug-2002  
C:Accession: A55567  
R:Tilstra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.  
Genomics 23, 480-485, 1994  
A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization  
A:Reference number: A55567; MUID:95137597; PMID:7835900  
A:Accession: A55567  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2871 <TIL>  
A:Cross-references: GB:L28748; NID:g508427; PIDN:AAA74122.1; PID:g508428  
C:Superfamily: fibrillin 1; EGF homology  
F:1201-1236/Domain: EGF homology <EGF>  
Query Match 12.6%; Score 368; DB 2; Length 2871;  
Best Local Similarity 31.5%; Pred. No. 2.4e-14;  
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;  
QY 244 DCSVNGGCEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNLCHEFCVNP-- 300  
DB 1200 ECSIMNGGCTCTNSGSEYECSCQPGFALMPDQRSC-----DIDE--CEDNPIC 1249  
QY 301 -----DQPGSYSCMCTGYRLAADOHRCEVDVDCILEPSPC--PQRCVNTQGGFECHCYP 353  
DB 1250 DGGQCTNIPGEYRCLCYDGFMA5MDKMTCLDVNECDLNENICLSGTCENTKGSFICHCDM 1309  
QY 354 NYDLVDGE--CVBPVDPG--FRANCEYQCOPLNQT--SYLCVCAEGFA-----PIP 398  
DB 1310 GYSKGGKGTCTD-INECEIGHNCDRHAVCTNTAGSKSCSPGWIGDKICTDLDECS 1368  
QY 399 HEHRCQMFNCTACPADCDPNTQAS--CECPGYILLDGGFTCTDIDE--EN----- 447  
DB 1369 NGTHMCSQH-----ADC--KNTMGSYRCLCKEGY-TGDFGFTCTDIDESENILGNGNQ 1419  
QY 448 -----GGF-----CS-----GVCHNLPCTFECICGPDSSALA 473  
DB 1420 CLNAPGGYRCECDMGFVPSADGKACEDIDECSLPNTICVFTGTCNNLPGLRCECEIGYELD 1479  
QY 474 RHIG-----TDCDSGK--VDGDSGSGEPPSPPTPGSTLTPPAVGLV 513  
DB 1480 RSGNGCTDNECLDPTTCTISGNCVNTPGSYTCDCPPD-----FELNPTRVGCV 1527  
RESULT 8  
A55184  
fibrillin-2 precursor - human  
N:Alternate names: protein DkrZp586A1519.1  
C:Species: Homo sapiens (man)  
C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 02-Aug-2002

C:Accession: A55184; T08744  
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.  
Genomics 22, 425-430, 1994  
A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene  
A:Reference number: A55184; MUID:95104855; PMID:7806230  
A:Accession: A55184  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1184 <ZHA>  
A:Cross-references: GB:X82494; NID:9575232; PIDN:CAA57876.1; PID:g575233  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z16471  
A:Accession: T08744  
A:Molecule type: mRNA  
A:Residues: 656-719, 'QDECLMGADHCSRROFCVNTLGSFYCVNHTVLCADGYILNAHRKCYD', 720-853, 'T', 855-1184  
A:Cross-references: EMBL:AL050095  
A:Experimental source: adult uterus; clone DKF2p586A1519  
C:Genetics:  
A:Gene: GDB:FBLN2  
A:Cross-references: GDB:293037; OMIM:135821  
A:Map position: 3p25-3p24  
A:Note: DKF2p586A1519.1  
C:Superfamily: fibulin-2; EGF homology  
C:Keywords: alternative splicing; extracellular matrix  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-1184/Product: fibulin-2 protein #status predicted <MAT>  
F:905-941/Domain: EGF homology <EGF>  
Query Match 12.6%; Score 367.5; DB 2; Length 1184;  
Best Local Similarity 30.0%; Pred. No. 1.2e-14;  
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;  
QY 154 EQOCEVKAD-GFLCEHFHPTC-----RPLAV-----EPGAAAAVSITYGTPFAARG 201  
Db 531 EGOSCESNPILGYPCN-HVMSLSCGEGLIVPEVRPPPEPAAAPRRVS-----EAE 582  
QY 202 ADFQALPVGSSAAV---APLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGCGEACNA 258  
Db 583 AGRALSIGTBAEPLNSLPGDDQDECLLPGEL-----COHLGIN 622  
QY 259 IPGAPRCQCPAGAAQADGRSC-----TASAT-----QS 287  
Db 623 TVGSYHCACFPFGLQDDGRTCRPEGHPPQPEAPQEPALKSEFSQVASTIPTLPPLQPN 682  
QY 288 CND--LCBHFVNPDPQGSVSCMCETGYRLAADQHRCEVDVDCILESPCP--QRCVNT 343  
Db 683 CKDNGPKQVC---STVGSAICSCFPGYATMADGVSCEDINECVTDLHTCSRGEHCYNT 739  
QY 344 QGGFECH---CYPNYDLVDGCEVPEVDPFCFRANCEYOCPLNQTSYLCV----- 389  
Db 740 LGSFHCYKALTCEPGYALKDGE-EDVDEC--AMGTHTCOP-----GFLCNTKGSFYCOA 792  
QY 390 ---CAEGFAPIPH-----EPHRCOMFCNOTACPADCDPNTQASCECPGY-I 432  
Db 793 RQRCHDGLQDPPEGNCVDINECTSLSEPCRPFGSCINTVGSYTCQRNPLI---CARGYHA 849  
QY 433 LDDGFICTDIDCENGFCSG---VCHNLPGTFECIC 466  
Db 850 SDDGAKCVDNBECEGVHRGEGVCHNLPGSYRCD 886  
RESULT 9  
A54105  
fibillin-2 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 02-Aug-2002  
C:Accession: A54105; S17063; S31101  
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecham, J. Cell Biol. 124, 855-863, 1994  
A:Title: Structure and expression of fibillin-2, a novel microfibrillar component prefered  
A:Reference number: A54105; MUID:94165150; PMID:8120105  
A:Accession: A54105

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-2918 <ZHA>  
A:Cross-references: GB:U03272  
R:Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17063  
A:Molecule type: mRNA  
A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>  
A:Cross-references: EMBL:X62009  
R:Milewicz, D.M.  
submitted to the EMBL Data Library, December 1992  
A:Reference number: S31101  
A:Accession: S31101  
A:Molecule type: mRNA  
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P',  
A:Cross-references: EMBL:X62009  
C:Genetics:  
A:Gene: GDB:FBN2  
A:Cross-references: GDB:128122; OMIM:121050  
A:Map position: 5q23-5q31  
C:Superfamily: fibillin 1; EGF homology  
C:Keywords: extracellular protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-2918/Product: fibillin-2 #status predicted <MAT>  
F:1245-1280/Domain: EGF homology <EGF>  
F:1970-2013/Domain: EGF homology <EGF>  
Query Match 12.6%; Score 366; DB 2; Length 2918;  
Best Local Similarity 33.1%; Pred. No. 3.2e-14;  
Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;  
QY 244 DCSVENGGCEHACNAIPGAPRCQCPAGAAQADGRSCATASATQSCNDLCEHFCVNPND-- 301  
Db 1244 EGMINGGCDTQCTNSEGSYECSCSEGYALMPDRSCA-----DIDE--CENNPDIC 1293  
QY 302 -----OPGSYSCMCETGYRLAADQHRCEVDVDCILESPCP--QRCVNTQGGFECHCYP 353  
Db 1294 DGGQCTNTPGEYRCLCYDGFMA5MDMKTCIDVNECDLNSNICMFGECENTKGSFICHCQL 1353  
QY 354 NYDLVDGE--CYEVPDPC--FRANCEYOCPLN-QTSYLCVCAEGFAPIPH-----EPHR 403  
Db 1354 GYSVAKGTGTGCTD-VDECEIGHNCDMASHCLNIPGSCFCSREGW--IGNGIKCIDLDE 1410  
QY 404 COMFCNOTACPADCDPNTQAS--CECPEGVILDDGFICTDIDE-----CENG----- 448  
Db 1411 CSNGTHQCSINAQC-VNTEGSRACSEGF-TGDGFTCSVDDECAENINLCENGQCLNVP 1468  
QY 449 -----GF-----CSGVCHNLPDGTFCICGPDLSALAHIG- 477  
Db 1469 GAYRCBCEMGFTPASDRSCQDIDEC5FQNICVSGTCNNLPGMHFHCICDDGYELDRGTGN 1528  
QY 478 -TDCD 481  
Db 1529 CTDDID 1533  
RESULT 10  
A47221  
fibillin 1 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jun-1995 #sequence\_revision 25-Apr-1997 #text\_change 02-Aug-2002  
C:Accession: A47221; S143355; S17064; S17062; S62111; A34198  
R:Corson, G.M.; Chalberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y. Genomics 17, 476-484, 1993  
A:Title: Fibillin binds calcium and is coded by cDNAs that reveal a multidomain str  
A:Reference number: A47221; MUID:94010947; PMID:7691719  
A:Accession: A47221  
A:Molecule type: mRNA  
A:Residues: 1-337, 'T', 339-1029 <COR>  
A:Cross-references: GB:X63556





R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: 223035  
A:Accession: T46488  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-741 <AAA>  
A:Cross-references: EMBL:AL137638  
A:Experimental source: adult testis; clone DKF2p43J065  
C:Genetics:  
A:Note: DKF2p43J065.1

Query Match 12.2%; Score 357; DB 2; Length 741;  
Best Local Similarity 31.6%; Pred. No. 3.6e-14;  
Matches 86; Conservative 40; Mismatches 102; Indels 44; Gaps 15;  
Qy 245 CSVENGGCEHACNAIPGAPRCQCPAGAAQADLRGSCA-SATQSCNDLCEHFCVPNPDPQ 303  
Db 68 CAMEHNCCEQLCVNVPFGVQCYSGVALAEDGRVAVDYCASENHGCEHCV-NAD-- 124  
Qy 304 GYSOMCETGYRLAADOHRCEVDVDCILEPSPQPCQRCVNTGGFECHCPYNDL-VDGEC 362  
Db 125 GSYLCQCEGFALNPDEKCTKIDYCASSNHGQCHECVNTDSDSYSCHLKGFTLNPKKT 184  
Qy 363 VEPVDPCE--RANCEYQCPQNTQSYLCVCAEGFAPTPH-----EPHRCQMF 408  
Db 185 CRRINYCALNPGCEHECVNNEE-SYCYRCHRGYTLDPNGKTCGRVDHCAQDHCCEQLC 243  
Qy 409 NOTACPADCDNTQAS--CECEGVILDDGF-ICTDDIDG--ENGFGSCVCHNLPTGF 462  
Db 244 -----LNTDSFVQCSEGLNEDLKTCGRVDYCLLSDHG--CEYSCVNMDSRF 291  
Qy 463 ECICGPDALARRHGTDCDSKVPD---GDG 491  
Db 292 ACQC-PEGHVLRSKGTC--AKLDSALGDRG 320

RESULT 13  
A49457  
fibulin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 02-Aug-2002  
C:Accession: A49457; S74095  
R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.  
J. Cell Biol. 123, 1269-1277, 1993  
A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with  
A:Reference number: A49457; MUID:94064787; PMID:8245130  
A:Accession: A49457  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1221 <PAN>  
A:Cross-references: GB:X75285; NID:g437046; PIDN:CAA53040.1; PID:g437047  
R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.  
Eur. J. Biochem. 240, 427-434, 1996  
A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix metalloproteinases  
A:Reference number: S74094; MUID:96439073; PMID:8841408  
A:Accession: S74095  
A:Molecule type: protein  
A:Residues: 236-238, 'x', 240-247; 260-275; 336-344, 'L', 346-361; 405-426; 566-568, 'EM', 569-589  
C:Superfamily: fibulin-2; EGF homology  
C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer  
F:1942-978/Domain: EGF homology <EGF>

Query Match 12.2%; Score 357; DB 2; Length 1221;  
Best Local Similarity 29.5%; Pred. No. 5.4e-14;  
Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;  
Qy 154 EQQCCEVAD-GFLCEFHPPATC-----RPLAV-----EPGAAAVS-----IT 192  
Db 521 EQQCESPNLGYPCN-HVMLSCCEBEPPLVPEVRRPPEAPRVSSEMAAREALS 579  
Qy 193 YGT-----PFAARGADFO--ALP-----VGSSAAVAPLGLQLM-----CTAPPG 229

Db 580 LGTEAELPNSLPGDDQDECLMLPGLCOHLICINTVGSYRCACFPGLQDGRTRCPDRG 639  
Qy 230 AVQGHWARE-APGAWDCSV-----ENGCGEACNAIPGAPRCQCPAGAA 272  
Db 640 APOLDTAREAPRSASQAQVSPNTIPLVPQPNCKDNGPCQVGVGDTAMSCFPGYA 699  
Qy 273 LQADGRSC-----TASATQSCNDLCEHFCVPNPDPQSYSC-----MCETGYRLAADOHR 322  
Db 700 IMADGVSCDEDDCLMGTDCS--WKQFCV---NTLGSFYCVNHTVLCAEGYILNA-HRK 753  
Qy 323 CEDVDDCILEPSPC--PORCVNTGGFECH---CYNYDLVDGECVEPVDPDFRA--NC 374  
Db 754 CVDINECVTLHTCTRAEHCNVTPGSCQYKALTCCEGYVLTGECTD-VDECVTGHNC 812  
Qy 375 E-YQCOPLNOTSYLCV---CAEGFAPTPH-----EPHRCQMFNQTACPA 415  
Db 813 QAGFSCON-TKGSFYCAQRCMDGLQDPGNCVDINECTSLLEPCRSFSCINTVGSY 871  
Qy 416 DCDPNTQASCECPGY-ILDDGFTCTDIDCEENGFCGSG---VCHNLPGTFECICGP 468  
Db 872 TCQRNPLV---CGRGYHANEESGCVDNBCECTGVHRCGEGOLCYNLPGSYRCDCKP 925

RESULT 14  
T09059  
notch4 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jan-2000  
C:Accession: T09059  
R:Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Loretz, C.;  
submitted to the EMBL Data Library, October 1997  
A:Description: Sequence of the mouse major histocompatibility locus class III region.  
A:Reference number: Z16543  
A:Accession: T09059  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1964 <ROW>  
A:Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564947  
C:Genetics:  
A:Gene: notch4  
A:Map position: 17  
A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1;  
1679/3; 1729/1; 1761/3  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
C:Keywords: receptor; signal transduction  
F:514-545/Domain: EGF homology <EGF>

Query Match 11.4%; Score 332.5; DB 2; Length 1964;  
Best Local Similarity 26.3%; Pred. No. 2.3e-12;  
Matches 128; Conservative 34; Mismatches 142; Indels 183; Gaps 31;  
Qy 134 GPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCE-FHFPTATCRPLAVEPGAAAAAVSIT 192  
Db 35 GGTCLRLSRGCGIC-----QC---APGLGETCQFPDPCR----- 66  
Qy 193 YGTFFAARGADFOAL---PVGSSAAVAPLGLQLMCTAPPG---AVQGHWAREAPGAWDC 245  
Db 67 -DTQLCKNGSCQALLPTPPSSRSPSTPLTPHFCTCPSGTCGDRQCQTHLEELCPPSF-C 124  
Qy 246 SVENGCGEACNAIPGAPRCQCPAGAAQADLRGSCA-SATQSCNDLCEHFCVPNPDPG- 304  
Db 125 S--NGG--HCYQASGRPQCCEPGWT-----GEOCQ-----LRDFCSANPCANGG 166

Qy 305 -----SYSCMCETGYRLAADOHRCE-DVDDCILEPSPCPQ---RCVNTGGFECHC--- 351  
Db 167 VCLATYPIQICRCPPGF-----EGHTCERDINECFLEPGCPQGTSCHTLGSYCLCPVG 222  
Qy 352 ---YNYDLVDGECVEPVDPDFRANCEYQCPPL---NNTSYLCVCAEGFAPTPHE----- 400  
Db 223 QEGPQCKLRKAGAC--PPGSCILNGG---TCQLVPEGHSTFHLCCLCPGPTGLDCEMNPDDC 277  
Qy 401 -PHRCQMFNQTACPADCDPNTQASCECPGYILDDGFICT-DIDECE-----NGGF 450

Db 278 VRHOCQ---NGATCLDGLDLYT---CLCPKTV---KWDCSEIDICEAARGPPRCRNGGT 328  
QY 451 C-----SGVCHNLPGTFECICGP----- 468  
Db 329 CONTAGSHVCVCSWVGAGCEENLDCAATCAPGTCIDRVGFSFCLCPGPGRTGLLCH 388  
QY 469 --DSALAR--HIGTDCDSGKVDG-----GDSGS-----GEPPPSP----- 499  
Db 389 LEDMCLSQPCHVNAQCSTNPLTGSTLCLCPGYSGSTCHQDLDECQMAQQGSPCEHGG 448  
QY 500 ---TPGS 503  
Db 449 CINTPGS 455

RESULT 15  
T43210  
fibulin-1D precursor - Caenorhabditis elegans (fragment)  
C:Species: Caenorhabditis elegans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T43210  
R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.  
submitted to the EMBL Data Library, June 1998  
A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character  
A:Reference number: Z22337  
A:Accession: T43210  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-589 <BAR>  
A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1  
C:Genetics:  
A:Note: intron positions not resolved (incomplete sequence)  
C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 11.4% Score 331; DB 2; Length 589;  
Best Local Similarity 28.1%; Pred. No. 1.1e-12; Indels 108; Gaps 28;  
Matches 126; Conservative 49; Mismatches 166

QY 103 GPLRGF-QWVTGDNNTSYSRWRLDNGAPLCPLCVASAAEATVPSEPI-----WEEQ 156  
Db 19 GCLRSFNKCCNGDIEITH--ASEIITGRPLNDPHVLHLGDRCASSHCHLCHDRGGEKV 75  
QY 157 QCEVKA-----DGFCEFHFP--ATCRPLAVE-----PGAAAAVSIYGTFFA--- 198  
Db 76 ECSSRGFDLAPDGMACVDHIDECATLMDCLSQRLNTPGSPFKCIRTLSCTGYAMDS 135  
QY 199 --ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQ-----GHWAREAPGAWDCS---- 246  
Db 136 ETERCRDVECNLGS--DCGPLYQCRNTQGSYRCDAKKCGDGLQNPMTGECTSITC 191  
QY 247 -----VENGGGE-----HACNA-----IPGAPRCQ-----CPAGAALQADGRSCTA 282  
Db 192 PNGYYPKNGMCNDIDECVTGHCNGAGEECVNTPGSFRCQKGNLCAHGYEVN----- 243  
QY 283 SATQSCNDL--CEH-----FCVNPDPQSGYSQWCEYVRLAADOHRCEVDVDCIL--- 331  
Db 244 GATGCEVDNECQQQVCGSMCEI---NLPGYKCKGPGFNFNDKAKCEDVDCEIKFAG 300  
QY 332 EPSPCQRCVNTQGGFECYCPNYDLV--DGECEVFPDPCFR--ANCEYQQQPLNQTSYLC 388  
Db 301 HVCDLASACINTIGSFECKKPGFQLASDGRRCEDVNECTTIGIAACEQKCVNI--PGSYQC 359  
QY 389 VCAEGFAIP-----HEPHRCQMF--NOTACPADCDPNTQAS--CECPGEG-ILDDGFI 438  
Db 360 ICDRGFALGPDGTCKEDIDECSIWAGSGNDLGMGC-INTKGSYLQCPFGYKIQPDGRT 418  
QY 439 CTDIDCEGNGGFCSG---VCHNLPGTFEC 464  
Db 419 CYDVECAMGE-CAGSDKVCVNTLGSFKC 446

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:20:02 ; Search time 10.5 Seconds  
(without alignments)  
2038.265 Million cell updates/sec

Title: US-09-509-994-2  
Perfect score: 2916  
Sequence: 1 MLGVILVGLALAGLGFAP.....PSPTPGSTLTTPPAGLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2916	100.0	575	1	TRBM_HUMAN	P07204 homo sapien
2	1903	65.3	577	1	TRBM_MOUSE	P15306 mus musculus
3	1095.5	37.6	356	1	TRBM_BOVIN	P06579 bos taurus
4	587	20.1	652	1	CD93_HUMAN	O89p33 homo sapien
5	547	18.8	644	1	CD93_MOUSE	O89p103 mus musculus
6	534.5	18.3	643	1	CD93_RAT	Q9et61 rattus norv
7	374	12.8	2907	1	FBN2_MOUSE	P61555 mus musculus
8	368	12.6	2871	1	FBN1_BOVIN	P98133 bos taurus
9	367.5	12.6	1184	1	FBF2_HUMAN	P98095 homo sapien
10	366	12.6	2871	1	FBN1_PIG	Q9tv36 sus scrofa
11	366	12.6	2911	1	FBN2_HUMAN	P35556 homo sapien
12	364	12.5	2871	1	FBN1_HUMAN	P35555 homo sapien
13	362	12.4	2871	1	FBN1_MOUSE	Q61554 mus musculus
14	357	12.2	1221	1	FBF2_MOUSE	P37889 mus musculus
15	356	12.2	956	1	MTN2_HUMAN	O00339 homo sapien
16	332.5	11.4	1964	1	NTC4_MOUSE	P31695 mus musculus
17	330.5	11.3	1712	1	LTB1_RAT	Q00918 rattus norv
18	326.5	11.2	956	1	MTN2_MOUSE	Q08746 mus musculus
19	324.5	11.1	712	1	FBF1_CAEEL	O77469 caenorhabdi
20	324.5	11.1	1394	1	LTBS_HUMAN	P22064 homo sapien
21	322.5	11.1	1595	1	LTBL_HUMAN	Q14766 homo sapien
22	321	11.0	2321	1	NTC3_HUMAN	Q9um47 homo sapien
23	315.5	10.8	443	1	FBF4_HUMAN	O95967 homo sapien
24	311	10.7	703	1	FBF1_HUMAN	P23142 homo sapien
25	308.5	10.6	443	1	FBF4_CRIGR	O55058 cricetus
26	308.5	10.6	2003	1	NTC4_HUMAN	Q9ubx5 homo sapien
27	308	10.6	448	1	FBF5_HUMAN	Q9ubx5 homo sapien
28	308	10.6	705	1	FBF1_MOUSE	Q08879 mus musculus
29	307	10.5	684	1	FBF1_CHICK	O73775 gallus gall
30	306.5	10.5	443	1	FBF4_MOUSE	Q9wvj9 mus musculus
31	302	10.4	448	1	FBF5_MOUSE	Q9wvh9 mus musculus
32	297	10.2	2531	1	NTC1_MOUSE	O01705 mus musculus
33	296	10.2	448	1	FBF5_RAT	Q9wvh8 rattus norv

RESULT 1

TRBM_HUMAN	294	10.1	2319	1	NTC3_RAT	Q9rl72 rattus norv
AC	292	10.0	2470	1	NTC2_MOUSE	Q35516 mus musculus
DT	289.5	9.9	493	1	FBF3_HUMAN	Q12805 homo sapien
DT	287	9.8	2471	1	NTC2_RAT	Q9w30 rattus norv
DT	285	9.8	2471	1	NTC2_HUMAN	Q04721 homo sapien
DT	284.5	9.8	2437	1	NTC1_BRARE	P46530 brachydanio
DT	280.5	9.6	2703	1	NOTC_DROME	P07207 drosophila
DT	280	9.6	2318	1	NTC3_MOUSE	Q61982 mus musculus
DT	278	9.5	2531	1	NTC1_RAT	Q07008 rattus norv
DT	276.5	9.5	493	1	FBF3_RAT	Q35568 rattus norv
DT	272	9.3	2556	1	NTC1_HUMAN	P46531 homo sapien
DT	269.5	9.2	1217	1	EGF_MOUSE	P01132 mus musculus

ALIGNMENTS

AC TRBM\_HUMAN STANDARD; PRT; 575 AA.  
P07204;  
01-APR-1988 (Rel. 07, Created)  
01-FEB-1991 (Rel. 17, Last sequence update)  
15-JUN-2002 (Rel. 41, Last annotation update)  
Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen).  
GN THBD OR THRM.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88004395; PubMed=2820710;  
RA Suzuki K., Kusumoto H., Devashiki Y., Nishioka J., Maruyama I.,  
RA Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.;  
RT "Structure and expression of human thrombomodulin, a thrombin  
RT receptor on endothelium acting as a cofactor for protein C  
RT activation.";  
RL EMBO J. 6:1891-1897(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88024950; PubMed=2822087;  
RA Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.;  
RT "Human thrombomodulin: complete cDNA sequence and chromosome  
RT localization of the gene.";  
RL Biochemistry 26:4350-4357(1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87317665; PubMed=2819876;  
RA Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.;  
RT "Human thrombomodulin gene is intron depleted: nucleic acid sequences  
RT of the cDNA and gene predict protein structure and suggest sites of  
RT regulatory control.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88227901; PubMed=2836377;  
RA Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,  
RA Devashiki Y., Maruyama I., Suzuki K.;  
RT "Gene structure of human thrombomodulin, a cofactor for thrombin-  
RT catalyzed activation of protein C.";  
RL J. Biochem. 103:281-285(1988).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carter C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Levasalho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce R.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann C.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20,";  
RL Nature 414:865-871(2001).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS.  
RX MEDLINE-94029900; PubMed-8216207;  
RA Gerleitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U.,  
RA Grinnell B.W.;  
RT "Identification of the predominant glycosaminoglycan-attachment site  
RT in soluble recombinant human thrombomodulin: potential regulation of  
RT functionality by glycosyltransferase competition for serine474,";  
RL Biochem. J. 295:131-140(1993).  
RN [7]  
RP STRUCTURE BY NMR OF 389-407.  
RX MEDLINE-96007474; PubMed-7559494;  
RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;  
RT "The structure of a 19-residue fragment from the C-loop of the fourth  
RT epidermal growth factor-like domain of thrombomodulin,";  
RL J. Biol. Chem. 270:23366-23372(1995).  
RN [8]  
RP STRUCTURE BY NMR OF 364-407.  
RX MEDLINE-96100636; PubMed-8528067;  
RA Meinelinger D.P., Hunter M.J., Komives E.A.;  
RT "Synthesis, activity, and preliminary structure of the fourth  
RT EGF-like domain of thrombomodulin,";  
RL Protein Sci. 4:1683-1695(1995).  
RN [9]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE-95034791; PubMed-7947766;  
RA Srinivasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;  
RT "Thrombin-bound structure of an EGF subdomain from human  
RT thrombomodulin determined by transferred nuclear Overhauser  
RT effects,";  
RL Biochemistry 33:13553-13560(1994).  
RN [10]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE-96276211; PubMed-8745396;  
RA Hrabal R., Komives E.A., Ni F.;  
RT "Structural resiliency of an EGF-like subdomain bound to its target  
RT protein, thrombin,"  
RL Protein Sci. 5:195-203(1996).  
RN [11]  
RP STRUCTURE BY NMR OF 405-444.  
RX MEDLINE-98035729; PubMed-9367781;  
RA Sampoli Benitez B.A., Hunter M.J., Meinelinger D.P., Komives E.A.;  
RT "Structure of the fifth EGF-like domain of thrombomodulin: an  
RT EGF-like domain with a novel disulfide-bonding pattern,";  
RL J. Mol. Biol. 273:913-926(1997).  
RN [12]  
RP VARIANT TED TYR-486.  
RX MEDLINE-95111115; PubMed-7811989;  
RA Oehlin A.-K., Marlar R.A.;  
RT "The first mutation identified in the thrombomodulin gene in a  
RT 45-year-old man presenting with thromboembolic disease,";  
RL Blood 85:330-336(1995).  
RN [13]  
RP VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.  
RX MEDLINE-97341986; PubMed-9198186;  
RA Oehlin A.-K., Norlund L., Marlar R.A.;  
RT "Thrombomodulin gene variations and thromboembolic disease,";  
RL Thromb. Haemost. 78:396-400(1997).  
RN [14]  
RP VARIANT VAL-473.  
RX MEDLINE-97206518; PubMed-9157575;  
RA Norlund L., Holm J., Zoller B., Oehlin A.-K.;  
RT "A common thrombomodulin amino acid dimorphism is associated with  
RT myocardial infarction,";  
RL Thromb. Haemost. 77:248-251(1997).  
RN [15]  
RP VARIANT THR-43.  
RX MEDLINE-99057299; PubMed-9843165;  
RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,  
RA Stubbs P.J., Manger Cats V., Ireland H.;  
RT "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,  
RT and the risk of myocardial infarction in men,";  
RL Thromb. Haemost. 80:743-748(1998).  
RN [16]  
RP VARIANT VAL-473.  
RX MEDLINE-21143723; PubMed-11245641;  
RA Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,  
RA Juneja H.;  
RT "Thrombomodulin Ala455Val polymorphism and risk of coronary heart  
RT disease,";  
RL Circulation 103:1386-1389(2001).  
CC -!- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR  
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS  
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE  
CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA  
CC SCISSIONS THE ACTIVATED COPACTORS OF THE COAGULATION MECHANISM,  
CC FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF  
CC THROMBIN GENERATED.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING  
CC THROMBOMODULIN.  
CC -!- POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED  
CC RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).  
CC -!- DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,  
CC ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE  
CC DEFECTS OF THE HAEMOPOIETIC SYSTEM WHICH CREATES A TENDENCY TO THE  
CC OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE  
CC PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.  
CC -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
CC -!- DATABASE: NAME-PROW; NOTE-CD guide CD141 entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd141.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; X05495; CAA29045.1; -;  
CC EMBL; M16552; AAB59508.1; -;  
CC EMBL; J02973; AAA61175.1; -;  
CC EMBL; D00210; BAA00149.1; -;  
CC EMBL; AL049651; CAB51954.1; -;  
CC PIR; A27073; A27073.  
CC PIR; A28307; A28307.  
CC PIR; A29680; A29680.  
CC PDB; 1EGT; 15-NOV-95.  
CC PDB; 1FGD; 20-JUN-96.  
CC PDB; 1FGE; 20-JUN-96.  
CC PDB; 1TMR; 08-JUN-95.  
CC PDB; 1ZAO; 29-JAN-96.  
CC PDB; 1ADX; 24-DEC-97.  
CC PDB; 2ADX; 24-DEC-97.  
CC GlycosuiteDB; P07204; -;  
CC Genew; HGNC:11784; THBD.

Query Match		100.08;	Score 2916;	DB 1;	Length 575;
Best Local Similarity		100.08;	Pred. No. 7.4e-185;		
Matches 516;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGVLVGLGALAGLGFAPAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM	60		
DB	1	MGVLVGLGALAGLGFAPAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM	60		
QY	61	TVRSSVAADVISILLNGDGVGRRRLWIGLQPLPGCGDPKRLGRLRGFWVTGDNNTSYS	120		
DB	61	TVRSSVAADVISILLNGDGVGRRRLWIGLQPLPGCGDPKRLGRLRGFWVTGDNNTSYS	120		
QY	121	RWRLDLNGLAPLGLPLCAVSAAEATVPSEPIWEQOQCEVKADGLCFBHFPPATCRPLAV	180		
DB	121	RWRLDLNGLAPLGLPLCAVSAAEATVPSEPIWEQOQCEVKADGLCFBHFPPATCRPLAV	180		
QY	181	EPGAAAAVSIYTGPFPAARGADFALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP	240		
DB	181	EPGAAAAVSIYTGPFPAARGADFALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP	240		
QY	241	GAWDCSVENGCGEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP	300		
DB	241	GAWDCSVENGCGEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP	300		
QY	301	DQPGSYSCMCTGYRLAADOHRCEVDYDCILEPSPCPORCVNTQGGFECHECPNTDLVDG	360		
DB	301	DQPGSYSCMCTGYRLAADOHRCEVDYDCILEPSPCPORCVNTQGGFECHECPNTDLVDG	360		
QY	361	ECVEPVDFCFRANCEYQCPQLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN	420		
DB	361	ECVEPVDFCFRANCEYQCPQLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN	420		
QY	421	TQASCECPGYILDDGFICTDIDECENGCGFCGVCHNLPGTFECICGPDALSARHIGTDC	480		
DB	421	TQASCECPGYILDDGFICTDIDECENGCGFCGVCHNLPGTFECICGPDALSARHIGTDC	480		
QY	481	DSGKVDGSDGSGEPPSPPTPGTSTLTPPAVGLVHSG	516		
DB	481	DSGKVDGSDGSGEPPSPPTPGTSTLTPPAVGLVHSG	516		
RESULT 2					
ID	TRBM_MOUSE	STANDARD;	PRT;	577	AA.
AC	P15306;				
DT	01-APR-1990 (Rel. 14, Created)				
DT	01-APR-1990 (Rel. 14, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Thrombomodulin precursor (Petomodulin) (TM).				
GN	THBD.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89128454; PubMed=2536925;				
RA	Dittman W.A., Majerus P.W.;				
RT	"Sequence of a cDNA for mouse thrombomodulin and comparison of the				
RT	predicted mouse and human amino acid sequences.";				
RL	Nucleic Acids Res. 17:802-802(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89008498; PubMed=2844823;				
RA	Dittman W.A., Kumada T., Sadler J.E., Majerus P.W.;				
RT	"The structure and function of mouse thrombomodulin. Phorbol				
RT	myristate acetate stimulates degradation and synthesis of				
RT	thrombomodulin without affecting mRNA levels in hemangioma cells.";				
RL	J. Biol. Chem. 263:15815-15822(1988).				
CC	-!- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR				
CC	THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS				
CC	COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE				

CC	ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA				
CC	SCISSIONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,				
CC	FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF				
CC	THROMBIN GENERATED.				
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING				
CC	THROMBOMODULIN (BY SIMILARITY).				
CC	-!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	EMBL; X14432; CAA32597.1; -				
DR	PIR; A32001; A32001.				
DR	PIR; S08488; S08488.				
DR	HSSP; P07204; LEGT.				
DR	MGI; MGI:98736; Thbd.				
DR	InterPro; IPR000152; Asx_hydroxyl.				
DR	InterPro; IPR000561; EGF-like.				
DR	InterPro; IPR001881; EGF_Ca.				
DR	InterPro; IPR001304; Lectin_C.				
DR	InterPro; IPR001491; Thrbomoduln.				
DR	Pfam; PF00008; EGF; 4.				
DR	Pfam; PF00059; Lectin_c; 1.				
DR	PRINTS; PR00907; THRBOMODULN.				
DR	SMART; SM00034; CLECT; 1.				
DR	SMART; SM00179; EGF_CA; 1.				
DR	SMART; SM00001; EGF_like; 5.				
DR	PROSITE; PS00010; ASX_HYDROXYL; 2.				
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.				
DR	PROSITE; PS01186; EGF_2; 3.				
DR	PROSITE; PS01187; EGF_CA; 2.				
DR	PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.				
KW	Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;				
KW	Glycoprotein; Signal; EGF-like domain.				
FT	SIGNAL	1	16	POTENTIAL.	
FT	CHAIN	17	577	THROMBOMODULIN.	
FT	DOMAIN	17	517	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	518	541	POTENTIAL.	
FT	DOMAIN	542	577	CYTOPLASMIC (POTENTIAL).	
FT	DOMAIN	240	280	EGF-LIKE 1.	
FT	DOMAIN	283	323	EGF-LIKE 2.	
FT	DOMAIN	324	362	EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).	
FT	DOMAIN	364	404	EGF-LIKE 4.	
FT	DOMAIN	403	439	EGF-LIKE 5.	
FT	DOMAIN	440	480	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).	
FT	DISULFID	244	255	BY SIMILARITY.	
FT	DISULFID	251	264	BY SIMILARITY.	
FT	DISULFID	266	279	BY SIMILARITY.	
FT	DISULFID	287	295	BY SIMILARITY.	
FT	DISULFID	291	307	BY SIMILARITY.	
FT	DISULFID	309	322	BY SIMILARITY.	
FT	DISULFID	328	339	BY SIMILARITY.	
FT	DISULFID	335	348	BY SIMILARITY.	
FT	DISULFID	350	361	BY SIMILARITY.	
FT	DISULFID	368	377	BY SIMILARITY.	
FT	DISULFID	373	387	BY SIMILARITY.	
FT	DISULFID	389	403	BY SIMILARITY.	
FT	DISULFID	407	416	BY SIMILARITY.	
FT	DISULFID	412	424	BY SIMILARITY.	
FT	DISULFID	426	438	BY SIMILARITY.	
FT	DISULFID	444	454	BY SIMILARITY.	
FT	DISULFID	449	463	BY SIMILARITY.	
FT	DISULFID	465	479	BY SIMILARITY.	
FT	CARBOHYD	113	143	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	243	243	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	256	256	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	408	408	N-LINKED (GLCNAC. . .) (POTENTIAL).	

ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA  
SCISSIONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,  
FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF  
THROMBIN GENERATED.  
-!- SUBCELLULAR LOCATION: Type I membrane protein.  
-!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING  
THROMBOMODULIN (BY SIMILARITY).  
-!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X14432; CAA32597.1; -  
PIR; A32001; A32001.  
PIR; S08488; S08488.  
HSSP; P07204; LEGT.  
MGI; MGI:98736; Thbd.  
InterPro; IPR000152; Asx\_hydroxyl.  
InterPro; IPR000561; EGF-like.  
InterPro; IPR001881; EGF\_Ca.  
InterPro; IPR001304; Lectin\_C.  
InterPro; IPR001491; Thrbomoduln.  
Pfam; PF00008; EGF; 4.  
Pfam; PF00059; Lectin\_c; 1.  
PRINTS; PR00907; THRBOMODULN.  
SMART; SM00034; CLECT; 1.  
SMART; SM00179; EGF\_CA; 1.  
SMART; SM00001; EGF\_like; 5.  
PROSITE; PS00010; ASX\_HYDROXYL; 2.  
PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
PROSITE; PS01186; EGF\_2; 3.  
PROSITE; PS01187; EGF\_CA; 2.  
PROSITE; PS00041; C\_TYPE\_LLECTIN\_2; 1.

Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;  
Glycoprotein; Signal; EGF-like domain.  
SIGNAL 1 16 POTENTIAL.  
CHAIN 17 577 THROMBOMODULIN.  
DOMAIN 17 517 EXTRACELLULAR (POTENTIAL).  
TRANSMEM 518 541 POTENTIAL.  
DOMAIN 542 577 CYTOPLASMIC (POTENTIAL).  
DOMAIN 240 280 EGF-LIKE 1.  
DOMAIN 283 323 EGF-LIKE 2.  
DOMAIN 324 362 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
DOMAIN 364 404 EGF-LIKE 4.  
DOMAIN 403 439 EGF-LIKE 5.  
DOMAIN 440 480 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
DISULFID 244 255 BY SIMILARITY.  
DISULFID 251 264 BY SIMILARITY.  
DISULFID 266 279 BY SIMILARITY.  
DISULFID 287 295 BY SIMILARITY.  
DISULFID 291 307 BY SIMILARITY.  
DISULFID 309 322 BY SIMILARITY.  
DISULFID 328 339 BY SIMILARITY.  
DISULFID 335 348 BY SIMILARITY.  
DISULFID 350 361 BY SIMILARITY.  
DISULFID 368 377 BY SIMILARITY.  
DISULFID 373 387 BY SIMILARITY.  
DISULFID 389 403 BY SIMILARITY.  
DISULFID 407 416 BY SIMILARITY.  
DISULFID 412 424 BY SIMILARITY.  
DISULFID 426 438 BY SIMILARITY.  
DISULFID 444 454 BY SIMILARITY.  
DISULFID 449 463 BY SIMILARITY.  
DISULFID 465 479 BY SIMILARITY.  
CARBOHYD 113 143 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 494 494 O-LINKED (GLYCOSAMINOGLYCAN) (BY
SQ SEQUENCE 577 AA; 61867 MW; B20E50B0FE745014 CRC64;
  SIMILARITY).
Query Match 65.3%; Score 1903; DB 1; Length 577;
Best Local Similarity 65.6%; Pred. No. 2.9e-118;
Matches 342; Conservative 46; Mismatches 125; Indels 8; Gaps 4;
QY 1 MGVVLGALAGLGPAPAPAPGPGGQCVHEPCFALPGPAPFALNASQICDGLRGLM 60
DB 1 MGVVLGALAGLGPAPAPGPGGQCVHEPCFALPGPAPFALNASQICDGLRGLM 60
QY 61 TVRSSVAADVLSLLNGDG-GVRRRLWIGLQPLPGGDKRLGLRGLFQWVTGDNNTSY 119
DB 61 TVRSSVAADVLSLLNGDG-GVRRRLWIGLQPLPGGDKRLGLRGLFQWVTGDNNTSY 119
QY 120 SHWRLDNGAPICGLCVANSAEAATVPSEPIWEEQCEVKADGFCFEPFATCRPLA 179
DB 120 SHWRLDNGAPICGLCVANSAEAATVPSEPIWEEQCEVKADGFCFEPFATCRPLA 179
QY 118 SWARPNDOTAPLCGLCVTVSTATEAAGPEAWEEKPCETETQGLCEFTFTASCRPLT 177
DB 118 SWARPNDOTAPLCGLCVTVSTATEAAGPEAWEEKPCETETQGLCEFTFTASCRPLT 177
QY 180 VEP-GAAAAAIVSYTGPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWARE 238
DB 180 VEP-GAAAAAIVSYTGPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWARE 238
QY 178 VNTDRPEAAHLSYTYNTPFGVSGADFTLPVGSAAVEPLGLELVCRAPGTSEGHWARE 237
DB 178 VNTDRPEAAHLSYTYNTPFGVSGADFTLPVGSAAVEPLGLELVCRAPGTSEGHWARE 237
QY 239 APGAWDCSVENGCGCHACNATPGAPROCPCAGALQADGRSCTASATOSCNDCHEFCVP 298
DB 239 APGAWDCSVENGCGCHACNATPGAPROCPCAGALQADGRSCTASATOSCNDCHEFCVP 298
QY 238 ATGAWNCVSVENGCGCHACNATPGAPROCPCAGALQADGRSCTASATOSCNDCHEFCVP 297
DB 238 ATGAWNCVSVENGCGCHACNATPGAPROCPCAGALQADGRSCTASATOSCNDCHEFCVP 297
QY 299 NPDPGYSVSCMCTGYRLAADOHRCEVDVDCILEPSPCPQRCVNTQGGFCHCYPNYDLY 358
DB 299 NPDPGYSVSCMCTGYRLAADOHRCEVDVDCILEPSPCPQRCVNTQGGFCHCYPNYDLY 358
QY 298 NAEVPSYSVSCMCTGYRLAADOHRCEVDVDCILEPSPCPQRCVNTQGGFCHCYPNYDLY 357
DB 298 NAEVPSYSVSCMCTGYRLAADOHRCEVDVDCILEPSPCPQRCVNTQGGFCHCYPNYDLY 357
QY 359 DGECEVPDPFRANCEYQCPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCD 418
DB 359 DGECEVPDPFRANCEYQCPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCD 418
QY 358 DGECEVELLDPCFSGNCEPCQSPFTDYRCICAGPAPKDPDEPHKCEMFCNETSCPADCD 417
DB 358 DGECEVELLDPCFSGNCEPCQSPFTDYRCICAGPAPKDPDEPHKCEMFCNETSCPADCD 417
QY 419 PNTQSCPCPGYILDDGICTDIDECENGCGFSGVCHNLPGTPECICGPDSSALARIHGT 478
DB 419 PNTQSCPCPGYILDDGICTDIDECENGCGFSGVCHNLPGTPECICGPDSSALARIHGT 478
QY 418 PNSTVPCPCPGYILDDGICTDIDECENGCGFSGVCHNLPGTPECICGPDSSALARIHGT 477
DB 418 PNSTVPCPCPGYILDDGICTDIDECENGCGFSGVCHNLPGTPECICGPDSSALARIHGT 477
QY 479 DCDGSKV---DGGSGSGEPSPPTGSLTTPPAVLVHSG 516
DB 479 DCDGSKV---DGGSGSGEPSPPTGSLTTPPAVLVHSG 516
QY 478 DCDPIPVREDATKEEGSGEPSPPTGSLTTPPAVLVHSG 518
DB 478 DCDPIPVREDATKEEGSGEPSPPTGSLTTPPAVLVHSG 518
RESULT 3
TRBM_BOVIN
AC P08579; STANDARD; PRT; 356 AA.
DT 01-JAN-1998 (Rel. 06, Created)
DT 01-JAN-1998 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thrombomodulin (Fetomodulin) (TM) (Fragment).
GN THBD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87067408; PubMed=3024152;
RA Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.;
RT "Characterization of a thrombomodulin cDNA reveals structural
RT similarity to the low density lipoprotein receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838(1986).
CC -! FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
CC SCAISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,
CC FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF
CC THROMBIN GENERATED.
```

```
CC -! SUBCELLULAR LOCATION: Type I membrane protein.
CC -! TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING
CC THROMBOMODULIN.
CC -! SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M14657; AAA30785.1; -.
CC PIR: A25918; A25918.
CC HSSP: P07204; 1TWR.
CC InterPro: IPR000152; ASX_HYDROXYL.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR001881; EGF_CA.
CC Pfam: PF00008; EGF_5.
CC SMART: SM00179; EGF_CA; 1.
CC SMART: SM00001; EGF_Like; 3.
CC PROSITE: PS00010; ASX_HYDROXYL; 2.
CC PROSITE: PS00022; EGF_1; FALSE_NEG.
CC PROSITE: PS01186; EGF_2; 3.
CC PROSITE: PS01187; EGF_CA; 2.
CC Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;
CC Glycoprotein; EGF-like domain.
CC NON_TER 1 1
CC DOMAIN 1 296 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 297 320 POTENTIAL.
CC DOMAIN 321 356 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 17 57 EGF-LIKE 1.
CC DOMAIN 60 98 EGF-LIKE 2.
CC DOMAIN 99 137 EGF-LIKE 3.
CC DOMAIN 139 179 EGF-LIKE 4.
CC DOMAIN 178 214 EGF-LIKE 5.
CC DOMAIN 215 254 EGF-LIKE 6.
CC DISULFID 21 32 BY SIMILARITY.
CC DISULFID 28 41 BY SIMILARITY.
CC DISULFID 43 56 BY SIMILARITY.
CC DISULFID 64 72 BY SIMILARITY.
CC DISULFID 68 82 BY SIMILARITY.
CC DISULFID 84 97 BY SIMILARITY.
CC DISULFID 103 114 BY SIMILARITY.
CC DISULFID 110 123 BY SIMILARITY.
CC DISULFID 125 136 BY SIMILARITY.
CC DISULFID 143 152 BY SIMILARITY.
CC DISULFID 148 162 BY SIMILARITY.
CC DISULFID 164 178 BY SIMILARITY.
CC DISULFID 182 191 BY SIMILARITY.
CC DISULFID 187 199 BY SIMILARITY.
CC DISULFID 201 213 BY SIMILARITY.
CC DISULFID 229 228 BY SIMILARITY.
CC DISULFID 224 237 BY SIMILARITY.
CC DISULFID 239 253 BY SIMILARITY.
CC CARBOHYD 271 271 O-LINKED (GLYCOSAMINOGLYCAN) (BY
CC SIMILARITY).
CC
CC SEQUENCE 356 AA; 37795 MW; 29B41F097ABE1093 CRC64;
Query Match 37.6%; Score 1095.5; DB 1; Length 356;
Best Local Similarity 63.4%; Pred. No. 2.2e-65;
Matches 189; Conservative 31; Mismatches 63; Indels 15; Gaps 6;
QY 229 GAVQGHWAREAPGAWDCSVENGCGCHACNATPGAPROCPCAGALQADGRSCTASATQSC 288
DB 5 GTEGEHRSAPGAWDCSVENGCGCHACNATPGAPROCPCAGALQADGRSCTASATQSC 64
QY 289 NDLCEHFCVNPDPGYSVSCMCTGYRLAADOHRCEVDVDCILEPSPCPQRCVNTQGGF 348
DB 65 HQLCEHFC--HLHGLGNYTCICEAGYQLAADOHRCEVDVDCILEPSPCPQRCVNTQGGF 122
QY 349 CHCYPNYDLVDGCEVPEVDPFRANCEYQCPLNOTSYLCVCAEGFAPIPHEPHRCOMFC 408
```

Db 123 CHCDTGVELVDGCEVDPVDFCDNCEYQCPVGRSEHKICAEGFAPVPGAPHKQCMFC 182

Qy 409 NOTACPADCPNTQASCECEGEGYLDGFTICTDIDCEGNGFCGSGVCHNLPGTFEICGP 468

Db 183 NOTSCPADCPHYTCRCPEGYIIDEGSTCTDINECDT-NICPGQCHNLPGTYEICGP 241

Qy 469 DSALARIHGTDCDSGV-----DGGSGSGEPPPTPGSTLTP-PA-VGLVHSG 516

Db 242 DSALSGQIGIDCDPTQVNEERGCTPDYG--GSGEPVSPPTPGATARSPAPAGPLHSG 297

RESULT 4

CD93\_HUMAN

ID CD93\_HUMAN STANDARD; PRT; 652 AA.

AC Q9NPY3; O00274;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Complement component C1q receptor precursor (Complement component 1, q subcomponent, receptor 1) (ClqR) (ClqR(p)) (Clq/MBL/SPA receptor) (CD93 antigen) (CDw93).

DE GN ClQRI OR CD93.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX PubMed=9047234;

RA Nepomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;

RA "cDNA cloning and primary structure analysis of ClqR(P), the human Clq/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.;"

RT Immunity 6:119-129(1997).

RL [2]

RN SEQUENCE FROM N.A., AND VARIANT ALA-318.

RP PubMed=11781389;

RX Steinberger P., Szekeres A., Wille S., Stockl J., Selenko N.,

RA Prager E., Staffler G., Madic O., Stockinger H., Knapp W.;

RA "Identification of human CD93 as the phagocytic Clq receptor (ClqR) by expression cloning.;"

RT J. Leukoc. Biol. 71:133-140(2002).

RL [3]

RN SEQUENCE FROM N.A.

RP MEDLINE=21638749; PubMed=11780052;

RX Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Levaslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., Lloyd K., McMurray A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.;"

RL Nature 414:865-871(2001).

RN [4]

RP SEQUENCE FROM N.A.

RX TISSUE=Leukocyte;

RA Straussberg R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RN [5]

RP CHARACTERIZATION.

RX PubMed=11994479;

RA McGreal E.P., Ikewaki N., Akatsu H., Morgan B.P., Gasque P.;

RA "Human ClqR is identical with CD93 and the mNI-11 antigen but does not bind Clq.;"

RT J. Immunol. 168:5222-5232(2002).

RL [6]

RN O-GLYCOSYLATION.

RP PubMed=10092817;

RX Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;

RA "ClqR is a heavily O-glycosylated cell surface protein involved in the regulation of phagocytic activity.;"

RL J. Immunol. 162:3583-3589(1999).

CC -1- FUNCTION: Receptor (or element of a larger receptor complex) for Clq, mannose-binding lectin (MBL2) and pulmonary surfactant protein A (SPA). May mediate the enhancement of phagocytosis in monocytes and macrophages upon interaction with soluble defense collagens. May play a role in intercellular adhesion.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: Highly expressed in endothelial cells, platelets, cells of myeloid origin, such as monocytes and neutrophils. Not expressed in cells of lymphoid origin.

CC -1- PTM: N- and O-glycosylated.

CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC -1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.

CC -1- CAUTION: Has been sometimes referred to as a collectin receptor.

CC -1- CAUTION: According to Ref.5, Clq is not a ligand for ClQRI.

CC -1- DATABASE: NAME=PROW; NOTE=PROW 3:1-6(2001);

CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/467246456\_g.htm".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U94333; AAB53110.1; -

DR EMBL; AL118508; CAC00597.1; -

DR EMBL; BC028075; AAB28075.1; -

DR HSSP; P35555; 1EMN.

DR MIM; 120577; -

DR InterPro; IPR000152; Asx\_hydroxyl.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR001304; Lectin\_C.

DR InterPro; IPR001187; Tissue\_factor.

DR Pfam; PF00008; EGF; 5.

DR Pfam; PF00059; lectin\_c; 1.

DR Pfam; PF01108; Tissue\_fac; 1.

DR SMART; SM00034; CLECT; 1.

DR SMART; SM00179; EGF\_CA; 3.

DR SMART; SM00001; EGF\_like; 2.

DR PROSITE; PS00010; ASX\_HYDROXYL; 3.

DR PROSITE; PS00615; C-TYPE\_LECTIN\_1; FALSE\_NEG.

DR PROSITE; PS00041; C-TYPE\_LECTIN\_2; 1.

DR PROSITE; PS01186; EGF\_2; 3.

DR PROSITE; PS01187; EGF\_CA; 3.

DR Repeat; EGF-like domain; Signal; Transmembrane; Glycoprotein; Receptor; Lectin; Polymorphism.

FT SIGNAL 1 21

FT CHAIN 22 652

FT DOMAIN 24 580

FT TRANSMEM 581 601

FT DOMAIN 602 652

FT DOMAIN 32 174

FT DOMAIN 260 301

FT COMPLEMENT COMPONENT C1Q RECEPTOR. EXTRACELLULAR (POTENTIAL).

FT POTENTIAL.

FT CYTOPLASMIC (POTENTIAL).

FT C-TYPE LECTIN.

FT EGF-LIKE 1.



FT DOMAIN 302 344 EGF-LIKE 2.  
 FT DOMAIN 345 384 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 385 426 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 427 468 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 594 601 POLY-LEU.  
 FT DISULFID 264 275 BY SIMILARITY.  
 FT DISULFID 271 285 BY SIMILARITY.  
 FT DISULFID 287 300 BY SIMILARITY.  
 FT DISULFID 306 317 BY SIMILARITY.  
 FT DISULFID 311 328 BY SIMILARITY.  
 FT DISULFID 330 343 BY SIMILARITY.  
 FT DISULFID 349 358 BY SIMILARITY.  
 FT DISULFID 354 367 BY SIMILARITY.  
 FT DISULFID 369 383 BY SIMILARITY.  
 FT DISULFID 389 400 BY SIMILARITY.  
 FT DISULFID 396 409 BY SIMILARITY.  
 FT DISULFID 411 425 BY SIMILARITY.  
 FT DISULFID 431 443 BY SIMILARITY.  
 FT DISULFID 439 452 BY SIMILARITY.  
 FT DISULFID 454 467 BY SIMILARITY.  
 FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 318 318 V -> A.  
 FT CONFLICT 22 22 /FTid=VAR\_013573  
 FT CONFLICT 36 36 T -> V (IN AA SEQUENCE).  
 FT CONFLICT 38 39 C -> T (IN AA SEQUENCE).  
 FT CONFLICT 155 155 TA -> RI (IN AA SEQUENCE).  
 FT CONFLICT 186 186 G -> N (IN REF. 1).  
 FT CONFLICT 492 492 S -> A (IN AA SEQUENCE).  
 FT CONFLICT 496 496 S -> Q (IN AA SEQUENCE).  
 FT CONFLICT 504 504 R -> G (IN AA SEQUENCE).  
 FT CONFLICT 541 541 P -> S (IN REF. 1).  
 SQ SEQUENCE 652 AA; 68560 MW; ECEAFCEAC5FCAC2 CRC64;

Query Match 20.1%; Score 587; DB 1; Length 652;  
 Best Local Similarity 31.9%; Pred. No. 1e-31;  
 Matches 182; Conservative 55; Mismatches 209; Indels 124; Gaps 32;

QY 2 LGVLVIGALALA--GLGPPAPAEPPQGGSCQVHDCFPALYPGPATFLNAAQICDGLRGHL 59  
 DB 5 MGLLLLLLLLLLTPQAGTGDTE----AVCVGTACTYAHSGKLSAAEAQHNCQNGNL 60  
 QY 60 MIVRSVAAD---VISILLNGDGVGR--RLATIGLQLPG-CGDPKRLGRLGFWMT 112  
 DB 61 AFVKSKEAQHVQRVLAQLLRALATARMSKFWIGLQREKGLDPSL--PLKGF5WVG 118  
 QY 113 GDNNTSYRWARLDLNGAPLCGLCA--VSAAEATVPSE-PIWEEQQC-----EVKAD 163  
 DB 119 GGEDTPYNWHKELRNSC--ISKRCVSLLDLSQPLPSRLPKWSEGPCSGSPGSGSNIE 176  
 QY 164 GPLCFEHPATCRPIAV-EPGAAAAVSIYVTGTPFAARGADFOALPYGVSSAAVAPLGLQL 222  
 DB 177 GFVCKEFGKMRPLALGGPG-----QVTYTPFTPTSSLEAVPFAAANVA----- 224  
 QY 223 MC-TAPPAVQGHW---AREAPGAWD-----CSVENGCEHAC-NAIRGAP 263  
 DB 225 -CGEGDKDETSHYFLCKEKAPDFDMGSSGPLCVSPRYKGFNFNGGCHQDFRGGDSF 283  
 QY 264 RQCPAGAAQADGRSCVATASQNDLCE--HFCVNPDPQPGSYSCMCTGYRLAADQH 321  
 DB 284 LGCRRPFRLLDLVTC-ASRNPSSSPCRGATCVLGP-HGKNYTCRCQGYOLDSSQL 341  
 QY 322 REEDVDDCILPSPCRVQNTQGGFECHVYPNYDLVDCGCEVPDPCFRANCYOCQPL 381  
 DB 342 DCVDVDEC--QDSPCAQECVNTPGGFCRCWVG-----EPGGP-----GEGACQDV 386  
 QY 382 NQTSYLCVCAEGFAPIPHEPHRCQMFNCQATACPADCPNTQAS--CECPGYIL--DDGF 437  
 DB 387 DB-----CALGRSP-----CAQGC-TNTDGSFHCSEEGYVLAGEDGT 423  
 QY 438 ICTDIDEC--ENGFCSGVCHNLRTFCICGPDPSALARHITGTDGSKV-----D 486  
 DB 424 QCQDVDECVGGPLCDLSLCFNTQGSFHCGLPLGWLAPN-GVSCWTGMPVSLGPPSPGPD 482

QY 487 GDSGSGE-----PPSPPTGSLTPPA 509  
 DB 483 EEDKGEKGVTPRAATASPTRGPECTPKA 512

RESULT 5  
 CD93\_MOUSE  
 ID CD93\_MOUSE STANDARD; PRT; 644 AA.  
 AC O89103;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Complement component clg receptor precursor (Complement component 1, q subcomponent, receptor 1) (ClqR) (ClqR(p)) (Clq/MBL/SPA receptor) (CD93 antigen) (Cell surface antigen AA4) (Lymphocyte antigen 68).  
 GN CLQRI OR CD93 OR CLQRP OR LY68 OR AA4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV;  
 RX PubMed=11074255;  
 RA Kim T.S., Park M., Nepomuceno R.R., Palmirini G., Winokur S., Cotman C.A., Bengtsson U., Tenner A.J.;  
 RT "Characterization of the murine homolog of ClqR(p): identical cellular expression pattern, chromosomal location and functional activity of the human and murine ClqR(p).";  
 RL Mol. Immunol. 37:377-389(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukemia;  
 RX PubMed=10403644;  
 RA Petrenko O., Beavlis A., Klaine M., Kittappa R., Godin I., Lemischka I.R.;  
 RT "The molecular characterization of the fetal stem cell marker AA4.";  
 RL Immunity 10:691-700(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV; TISSUE=Spleen, and Endothelial cells;  
 RX MEDLINE=99359842; PubMed=10430665;  
 RA Norsworthy P.J., Taylor P.R., Walport M.J., Botto M.;  
 RT "Cloning of the mouse homolog of the 126-kDa human Clq/MBL/SP-A receptor, ClqRp.";  
 RL Mamm. Genome 10:789-793(1999).  
 CC -I- FUNCTION: Receptor (or element of a larger receptor complex) for Clq, mannose-binding lectin (MBL2) and pulmonary surfactant protein A (SPA). May mediate the enhancement of phagocytosis in monocytes and macrophages upon interaction with soluble defense collagens. May play a role in intercellular adhesion. Marker for early multipotent hematopoietic precursor cells. May play a role in cell-cell interactions during hematopoietic and vascular development.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- TISSUE SPECIFICITY: Expressed in lung, heart and bone marrow. Expressed at lower level in ovary, whole embryo and fetal liver. Not detected in brain, adult liver or thymus. Highly expressed in peritoneal cavity and bone marrow macrophages. Not detected in epithelial cells.  
 CC -I- DEVELOPMENTAL STAGE: First detectable in day 9 embryos, in the endocardium and vascular endothelium in the anterior part of the embryo. Expression in endothelial cells, initially restricted to aorta, omphalomesenteric and umbilical arteries, later extends to subcardinal veins, intersomitic arteries and perineural vessels. On day 10, detectable in the entire embryo.  
 CC -I- PTM: N- and O-glycosylated (By similarity).  
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -I- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -



the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF074856; AAC63274.1; -  
EMBL; AF081789; AAC62649.1; -  
EMBL; AF099939; AAD47906.1; -  
EMBL; AF099938; AAD47906.1; JOINED.  
HSP; P35555; 1EMN.  
MGD; MGI:106664; Ly68.  
InterPro: IPR000152; Asx\_hydroxyl.  
InterPro: IPR000561; EGF-like.  
InterPro: IPR001881; EGF\_Ca.  
InterPro: IPR001304; Lectin\_C.  
Pfam: PF00008; EGF; 5.  
Pfam: PF00059; lectin\_c; 1.  
SMART; SM00034; CLECT; 1.  
SMART; SM00179; EGF\_CA; 3.  
SMART; SM00001; EGF\_like; 2.  
PROSITE; PS00010; ASX\_HYDROXYL; 3.  
PROSITE; PS00615; C\_TYPE\_LECTIN\_1; FALSE\_NEG.  
PROSITE; PS0041; C\_TYPE\_LECTIN\_2; 1.  
PROSITE; PS01186; EGF\_2; 3.  
PROSITE; PS01187; EGF\_CA; 3.  
Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;  
Repeat; Lectin.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 644 COMPLEMENT COMPONENT C1Q RECEPTOR.  
FT DOMAIN 23 572 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 573 593 POTENTIAL.  
FT DOMAIN 594 644 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 31 173 C-TYPE LECTIN.  
FT DOMAIN 257 298 EGF-LIKE 1.  
FT DOMAIN 299 341 EGF-LIKE 2.  
FT DOMAIN 342 381 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 382 423 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 424 465 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
FT DISULFID 261 272 BY SIMILARITY.  
FT DISULFID 268 282 BY SIMILARITY.  
FT DISULFID 284 297 BY SIMILARITY.  
FT DISULFID 303 314 BY SIMILARITY.  
FT DISULFID 308 325 BY SIMILARITY.  
FT DISULFID 327 340 BY SIMILARITY.  
FT DISULFID 346 355 BY SIMILARITY.  
FT DISULFID 351 364 BY SIMILARITY.  
FT DISULFID 366 380 BY SIMILARITY.  
FT DISULFID 386 397 BY SIMILARITY.  
FT DISULFID 393 406 BY SIMILARITY.  
FT DISULFID 408 422 BY SIMILARITY.  
FT DISULFID 428 440 BY SIMILARITY.  
FT DISULFID 436 449 BY SIMILARITY.  
FT DISULFID 451 464 BY SIMILARITY.  
FT CARBOHYD 102 102 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 322 322 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT SEQUENCE 644 AA; 69354 MW; EB4351648BF8635A CRC64;  
Query Match 18.8%; Score 547; DB 1; Length 644;  
Best Local Similarity 30.28; Pred. No. 4, 3e-29;  
Matches 168; Conservative 55; Mismatches 210; Indels 124; Gaps 27;  
5 LVLAGALAGLGFAPAPBQPGSQCVHDCFALYPGPATFLNASQICDGLRGLMTVRS 64  
9 LLLGLLGPWAG--AAADSQ--AVVCEGTACTAHWKLISAAEAQHRCNENGSLATVKS 64  
65 SVAA----DVISLLNGDGGVGR--RLWIGLQLPPGCGDPRRLGRLGFQVWTGDNWTS 118  
65 EEEARHVQALQLKLTAPLEAKMKFWIGLQREKGNCTYHDL-PMRGFSWVGGEATFA 123  
119 YSRWRLDNLGAPICGCLCAVSAEAATVPSE-PTWEEQQCEV-----KADGFLCEHF 171  
124 YSNWYKASKSCIFKRCVSLILDLSLTPHPSHLPKWHSPCGTPEAPGNSIEGFLCKFNP 183

QY 172 PATCRPLAV-EPGAAAAAVSITYGTFFAARGADFOALPVGSSAAVAPLGLQ-----LM 223  
DB 184 KGMRPLALGPG-----RVITYTPQATTSLEAVPFASVANVA-CGDEAKSETHYFL 236  
QY 224 CT-APPGAVOGHWAREAP-----GAWDCSVENGCGEHAC-NAIPGAPRCQCAGAAALQADG 277  
DB 237 CNEKTPGIF--HWGSSGPLCVSPKFGCSFNNGCQDQCFEGGDSFGRCGRPFRLDDL 294  
QY 278 RSTASATQSCNDLCEHFCVNPNDPG-----SYSCMETGYRLAADOHRCEVDV 327  
DB 295 VTCAS-----RNPCSSNPCTGGGMCHSVPLSENITCRCPGSLQDSSQVHCVDID 344  
QY 328 DCILSPSPORCVNTQGGFECHCPYNYLDVG---ECVEPVDPCFRANCEVCQCPQ 384  
DB 345 EC--QDSPCAQDCVNTLGSFHCCEWVGQ--PSGKEEACEDVDECAANSP----- 392  
QY 385 SYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQAS--CECEGYIL--DDGFICT 440  
DB 393 -----CAQGGCI-----NTDGSFYCSCKEGIYVSGEDSTQCE 423  
QY 441 DIDECE--GGFCSGVCHNLPGTFECICGPDPSALARHGTDCDSG-----KVD 486  
DB 424 DIDECDARGNPDCLSCFTDGSFRGCGPPGWELAPN-GVFCRSCTVSELPARPPQKED 482  
QY 487 GDSGSGEPSPPTPGS 503  
DB 483 NDRKESTMPPTPESS 499  
RESULT 6  
CD93\_RAT  
ID CD93\_RAT STANDARD; PRT; 643 AA.  
AC O9FT61; O9JI26;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Complement component C1q receptor precursor (Complement component 1, q subcomponent, receptor 1) (C1qR) (C1qR(p)) (C1q/MBL/SPA receptor)  
DE (CD93 antigen), receptor 1 (Cell surface antigen AA4).  
DE C1QR1 OR CD93 OR C1QRP.  
GN Rattus norvegicus (Rat).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PVG; TISSUE-Natural killer cells;  
RX MEDLINE=20545218; PubMed=11093152;  
RA Lovik G., Vaage J.T., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;  
RT "Characterization and molecular cloning of rat C1qR, a receptor on NK cells.";  
RL Eur. J. Immunol. 30:3355-3362(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Wistar; TISSUE-Lung;  
RX MEDLINE=20507883; PubMed=10934210;  
RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;  
RT "Molecular and cellular properties of the rat AA4 antigen, a C-type lectin-like receptor with structural homology to thrombomodulin.";  
RL J. Biol. Chem. 275:34382-34392(2000).  
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for C1q, mannose-binding lectin (MBL2) and pulmonary surfactant protein A (SPA). May mediate the enhancement of phagocytosis in monocytes and macrophages upon interaction with soluble defense collagens. May play a role in intercellular adhesion.  
CC SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and heart. Expressed at lower level in brain, thymus, liver, spleen, intestine, kidney, adrenal gland, muscle and testis. Expressed on endothelial cells, platelets, undifferentiated monocytes and circulating natural killer cells.  
CC -!- PTM: N- and O-glycosylated (By similarity).

CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL; AF136537; AAG01572.1; -;  
 DR EMBL; AF160978; AAF80402.1; -;  
 DR HSP; P35555; 1EMN.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001304; Lectin\_C.  
 DR Pfam; PF00008; EGF; 5.  
 DR Pfam; PF00059; lectin\_c; 1.  
 DR SMART; SM00034; CLECT; 1.  
 DR SMART; SM00181; EGF; 5.  
 DR SMART; SM00179; EGF\_Ca; 5.  
 DR SMART; SM00001; EGF\_like; 2.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 3.  
 DR PROSITE; PS00615; C-TYPE LECTIN\_1; FALSE\_NEG.  
 DR PROSITE; PS00041; C-TYPE LECTIN\_2; 1.  
 DR PROSITE; PS01186; EGF\_2; 3.  
 DR PROSITE; PS01187; EGF\_CA; 3.  
 KW Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;  
 FT Repeat; Lectin.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 643 COMPLEMENT COMPONENT C1Q RECEPTOR.  
 FT DOMAIN 24 571 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 572 592 POTENTIAL.  
 FT DOMAIN 593 643 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 31 173 C-TYPE LECTIN.  
 FT DOMAIN 257 298 EGF-LIKE 1.  
 FT DOMAIN 299 341 EGF-LIKE 2.  
 FT DOMAIN 342 381 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 382 423 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 424 463 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 FT DISULFID 261 272 BY SIMILARITY.  
 FT DISULFID 268 282 BY SIMILARITY.  
 FT DISULFID 284 297 BY SIMILARITY.  
 FT DISULFID 303 314 BY SIMILARITY.  
 FT DISULFID 308 325 BY SIMILARITY.  
 FT DISULFID 327 340 BY SIMILARITY.  
 FT DISULFID 346 355 BY SIMILARITY.  
 FT DISULFID 351 364 BY SIMILARITY.  
 FT DISULFID 366 380 BY SIMILARITY.  
 FT DISULFID 386 397 BY SIMILARITY.  
 FT DISULFID 393 406 BY SIMILARITY.  
 FT DISULFID 408 422 BY SIMILARITY.  
 FT DISULFID 428 437 BY SIMILARITY.  
 FT DISULFID 433 446 BY SIMILARITY.  
 FT DISULFID 448 461 BY SIMILARITY.  
 FT CARBOHYD 322 322 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 498 498 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CONFLICT 417 417 E -> K (IN REF. 2).  
 SQ SEQUENCE 643 AA; 68781 MW; 9AE4C33AD943DB6 CRC64;  
 Query Match 18.3%; Score 534.5; DB 1; Length 643;  
 Best Local Similarity 29.9%; Pred. No. 2.8e-28;  
 Matches 167; Conservative 56; Mismatches 208; Indels 127; Gaps 28;  
 QY 4 VLVLGALAGLGFPAPEQPGSQVCHDFALYFGPATFLNASQICDGLRHLMTVR 63  
 Db LLLGLLQLWAGAANDSE---AVCEGTACTYTAHMKLSAAEAQHRCHNENGNGNLATVK 63  
 QY 64 SSVA---DVISLLN---GDGVGRRRLWIGLQIPPGCGDKRGLPLGFWQVTDN 115  
 Db SEEAEHVRQEAQLLTKKAPSETKIG--KFWIGLQREKRGKTYHDL-PMKGFSSWVGGE 120

QY 116 NTSYSRWARDLNG--APLGGPLCVAVSAAETVPSE-PIWEEQOCEV-----KADGFL 166  
 Db DTITSNWKASKSSCKRCVSLDLKPH--PSHLKPKWHESPCGTDPAFGNSTEGFL 178  
 QY 167 CEFHFFATCRPLAV-BPGAAAAAVSTYGTFFAARCAADFOALPVGSSAAVAPLGLQ---- 221  
 Db CKNFKMGCSPLALGGPG-----QUTYTPQATITSSLKAVPFASVANV-CGDAESK 231  
 QY 222 ---LMCTAPPAGVQGHAREAP-----GAWDCSVENGCEHAC-NAITPGAPRCQCPAGAAL 273  
 Db TNYYLCKETTAGV-FHWGSSGPLCVSPKFGCSFNNGCCQDCFEFGDGFRCGRPGRL 290  
 QY 274 QADGRSCTASATOSCNLDCEHFCVNPDPQG-----SYSCMCTGYRLAADOHRC 323  
 Db LDDLVTAS-----RNPCCSNPCTGGGMCHSVPLSENYTCHCPRGYQLDSSQVHC 340  
 QY 324 EDVDDCILPEPSPCPQCVNTQGGFECHCYPNYDLVDG--ECVEPVDPFCFRANCEYQCPL 381  
 Db VIDEDEC--EDSPCDQECINTPGGFHCECHWGVQSSSKSEACEDVDEC----- 386  
 QY 382 NQTSYLCVCAEGFAPIPHEPHRCMFCNOTACPADCDPNTQAS--CECEGYIL--DDGF 437  
 Db --TAAYSPCAQG-----CT-----NTDGSFYCSCKEYIMSGEDST 420  
 QY 438 ICTDIDCEENGFCPSGVCHNLPGTFECICGPDASALAHIGTDCDSG-----KV 485  
 Db QCEIDDEC-IGNPCDTLICINTDGSFRGCPAGFELAPN-GVSCTRGSMFSELPARPPQKE 478  
 QY 486 DGGDGSSEPPSPPTPGS 503  
 Db DKGDGKSTVPLTEMPGS 496  
 RESULT 7  
 FBN2\_MOUSE  
 ID FBN2\_MOUSE STANDARD; PRT; 2907 AA.  
 AC O61555; O61557;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Fibrillin 2 precursor.  
 GN FBN2 OR FBN-2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95263670; PubMed=7744963;  
 RA Zhang H., Hu W., Ramirez F.;  
 RT "Developmental expression of fibrillin genes suggests heterogeneity  
 RT of extracellular microfibrils.";  
 RL J. Cell Biol. 129:1165-1176(1995).  
 RN [2]  
 RP SEQUENCE OF 210-317 FROM N.A.  
 RX MEDLINE=94140368; PubMed=8307578;  
 RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,  
 RA Francke U.;  
 RT "Fibrillin genes map to regions of conserved mouse/human synteny on  
 RT mouse chromosomes 2 and 18.";  
 RL Genomics 18:667-672(1993).  
 CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
 CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
 CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
 CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL; L39790; AAA74908.1; "	DOMAIN	2484	2524	EGF-LIKE 42, CALCIUM-BINDING.
CC	EMBL; S69359; AAC60685.1; "	DOMAIN	2525	2563	EGF-LIKE 43, CALCIUM-BINDING.
CC	HSSP; P35555; LEW.	DOMAIN	2564	2606	EGF-LIKE 44, CALCIUM-BINDING.
DR	MGI:95490; Fbn2.	DOMAIN	2607	2646	EGF-LIKE 45, CALCIUM-BINDING.
DR	InterPro; IPR000152; Asx_hydroxyl.	DOMAIN	2647	2687	EGF-LIKE 46, CALCIUM-BINDING.
DR	InterPro; IPR000561; EGF-like.	DOMAIN	2688	2727	EGF-LIKE 47, CALCIUM-BINDING.
DR	InterPro; IPR001881; EGF-Ca.	DISULFID	115	124	BY SIMILARITY.
DR	InterPro; IPR001438; EGF-II.	DISULFID	119	130	BY SIMILARITY.
DR	InterPro; IPR002212; Fibril-assoc.	DISULFID	132	141	BY SIMILARITY.
DR	Pfam; PF00008; EGF; 46.	DISULFID	143	159	BY SIMILARITY.
DR	Pfam; PF00683; Tb; 9.	DISULFID	153	164	BY SIMILARITY.
DR	PRINTS; PR00010; EGFBLLOOD.	DISULFID	166	175	BY SIMILARITY.
DR	SMART; SM00179; EGF_CA; 43.	DISULFID	180	190	BY SIMILARITY.
DR	SMART; SM00001; EGF_like; 3.	DISULFID	184	196	BY SIMILARITY.
DR	PROSITE; PS00010; ASX_HYDROXYL; 43.	DISULFID	198	207	BY SIMILARITY.
DR	PROSITE; PS00022; EGF_1; 2.	DISULFID	207	292	BY SIMILARITY.
DR	PROSITE; PS01186; EGF_2; 36.	DISULFID	287	301	BY SIMILARITY.
DR	PROSITE; PS01187; EGF_CA; 43.	DISULFID	303	316	BY SIMILARITY.
DR	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;	DISULFID	322	334	BY SIMILARITY.
KW	Repeat; Signal; Multigene family.	DISULFID	329	343	BY SIMILARITY.
FT	SIGNAL 1 28	DISULFID	345	358	BY SIMILARITY.
FT	CHAIN 29 2907	DISULFID	491	503	BY SIMILARITY.
FT	DOMAIN 111 142	DISULFID	498	512	BY SIMILARITY.
FT	DOMAIN 145 176	DISULFID	514	526	BY SIMILARITY.
FT	DOMAIN 176 208	DISULFID	532	542	BY SIMILARITY.
FT	DOMAIN 276 317	DISULFID	537	551	BY SIMILARITY.
FT	DOMAIN 318 359	DISULFID	553	566	BY SIMILARITY.
FT	REPEAT 360 426	DISULFID	572	584	BY SIMILARITY.
FT	DOMAIN 487 527	DISULFID	579	593	BY SIMILARITY.
FT	DOMAIN 528 567	DISULFID	595	608	BY SIMILARITY.
FT	DOMAIN 568 609	DISULFID	614	625	BY SIMILARITY.
FT	DOMAIN 610 650	DISULFID	620	634	BY SIMILARITY.
FT	DOMAIN 651 691	DISULFID	636	649	BY SIMILARITY.
FT	REPEAT 692 760	DISULFID	655	666	BY SIMILARITY.
FT	DOMAIN 761 802	DISULFID	661	675	BY SIMILARITY.
FT	DOMAIN 803 844	DISULFID	677	690	BY SIMILARITY.
FT	DOMAIN 845 883	DISULFID	765	777	BY SIMILARITY.
FT	DOMAIN 948 989	DISULFID	772	786	BY SIMILARITY.
FT	REPEAT 990 1065	DISULFID	788	801	BY SIMILARITY.
FT	DOMAIN 1066 1107	DISULFID	801	819	BY SIMILARITY.
FT	DOMAIN 1108 1150	DISULFID	814	828	BY SIMILARITY.
FT	DOMAIN 1151 1192	DISULFID	830	843	BY SIMILARITY.
FT	DOMAIN 1193 1234	DISULFID	849	859	BY SIMILARITY.
FT	DOMAIN 1235 1275	DISULFID	854	868	BY SIMILARITY.
FT	DOMAIN 1276 1317	DISULFID	870	883	BY SIMILARITY.
FT	DOMAIN 1318 1359	DISULFID	952	964	BY SIMILARITY.
FT	DOMAIN 1360 1400	DISULFID	959	973	BY SIMILARITY.
FT	DOMAIN 1401 1441	DISULFID	975	988	BY SIMILARITY.
FT	DOMAIN 1442 1483	DISULFID	1070	1082	BY SIMILARITY.
FT	DOMAIN 1484 1524	DISULFID	1077	1091	BY SIMILARITY.
FT	DOMAIN 1525 1565	DISULFID	1093	1106	BY SIMILARITY.
FT	REPEAT 1566 1642	DISULFID	1112	1124	BY SIMILARITY.
FT	DOMAIN 1643 1684	DISULFID	1119	1133	BY SIMILARITY.
FT	DOMAIN 1685 1726	DISULFID	1135	1149	BY SIMILARITY.
FT	REPEAT 1727 1800	DISULFID	1155	1167	BY SIMILARITY.
FT	DOMAIN 1801 1842	DISULFID	1162	1176	BY SIMILARITY.
FT	DOMAIN 1843 1884	DISULFID	1178	1191	BY SIMILARITY.
FT	DOMAIN 1885 1926	DISULFID	1197	1209	BY SIMILARITY.
FT	DOMAIN 1927 1965	DISULFID	1204	1218	BY SIMILARITY.
FT	DOMAIN 1966 2008	DISULFID	1220	1233	BY SIMILARITY.
FT	DOMAIN 2009 2048	DISULFID	1239	1250	BY SIMILARITY.
FT	DOMAIN 2049 2090	DISULFID	1246	1259	BY SIMILARITY.
FT	REPEAT 2091 2163	DISULFID	1261	1274	BY SIMILARITY.
FT	DOMAIN 2164 2205	DISULFID	1280	1292	BY SIMILARITY.
FT	DOMAIN 2206 2245	DISULFID	1287	1301	BY SIMILARITY.
FT	DOMAIN 2246 2286	DISULFID	1303	1316	BY SIMILARITY.
FT	DOMAIN 2287 2330	DISULFID	1322	1334	BY SIMILARITY.
FT	DOMAIN 2331 2372	DISULFID	1329	1343	BY SIMILARITY.
FT	REPEAT 2373 2441	DISULFID	1345	1358	BY SIMILARITY.
FT	DOMAIN 2442 2483	DISULFID	1364	1377	BY SIMILARITY.
FT		DISULFID	1371	1386	BY SIMILARITY.
FT		DISULFID	1388	1399	BY SIMILARITY.
FT		DISULFID	1405	1418	BY SIMILARITY.







DT	16-OCT-2001 (Rel. 40, Created)	FT	DOMAIN	1028	1069	EGF-LIKE 15, CALCIUM-BINDING.
DT	16-OCT-2001 (Rel. 40, Last sequence update)	FT	DOMAIN	1070	1112	EGF-LIKE 16, CALCIUM-BINDING.
DT	15-JUN-2002 (Rel. 41, last annotation update)	FT	DOMAIN	1113	1154	EGF-LIKE 17, CALCIUM-BINDING.
DE	Fibrillin 1 precursor.	FT	DOMAIN	1155	1196	EGF-LIKE 18, CALCIUM-BINDING.
GN	FN1.	FT	DOMAIN	1197	1237	EGF-LIKE 19, CALCIUM-BINDING.
OS	Sus scrofa (pig).	FT	DOMAIN	1238	1279	EGF-LIKE 20, CALCIUM-BINDING.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	FT	DOMAIN	1280	1321	EGF-LIKE 21, CALCIUM-BINDING.
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.	FT	DOMAIN	1322	1362	EGF-LIKE 22, CALCIUM-BINDING.
OX	NCBI_TaxID=9823;	FT	DOMAIN	1363	1403	EGF-LIKE 23, CALCIUM-BINDING.
RP	[1]	FT	DOMAIN	1404	1445	EGF-LIKE 24, CALCIUM-BINDING.
RP	SEQUENCE FROM N.A.	FT	DOMAIN	1446	1486	EGF-LIKE 25, CALCIUM-BINDING.
RC	TISSUE=Lung;	FT	DOMAIN	1487	1527	EGF-LIKE 26, CALCIUM-BINDING.
RC	MEDLINE=99156858; PubMed=10036187;	FT	REPEAT	1528	1605	TGFBP 4.
RA	Biery N.J.; Eldadah Z.A.; Moore C.S.; Stetten G.; Spencer F.,	FT	DOMAIN	1606	1647	EGF-LIKE 27, CALCIUM-BINDING.
RA	Dietz H.C.;	FT	DOMAIN	1648	1688	EGF-LIKE 28, CALCIUM-BINDING.
FT	"Revised genomic organization of FN1 and significance for regulated	FT	REPEAT	1689	1765	TGFBP 5.
FT	gene expression."	FT	DOMAIN	1766	1807	EGF-LIKE 29, CALCIUM-BINDING.
RL	Genomics 56:70-77(1999).	FT	DOMAIN	1808	1848	EGF-LIKE 30, CALCIUM-BINDING.
CC	-/- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS	FT	DOMAIN	1849	1890	EGF-LIKE 31, CALCIUM-BINDING.
CC	THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE	FT	DOMAIN	1891	1929	EGF-LIKE 32, CALCIUM-BINDING.
CC	LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.	FT	DOMAIN	1930	1972	EGF-LIKE 33, CALCIUM-BINDING.
CC	-/- FTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER	FT	DOMAIN	1973	2012	EGF-LIKE 34, CALCIUM-BINDING.
CC	FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE	FT	DOMAIN	2013	2054	EGF-LIKE 35, CALCIUM-BINDING.
CC	MICROFIBRILS.	FT	REPEAT	2055	2126	TGFBP 6.
CC	-/- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.	FT	DOMAIN	2127	2165	EGF-LIKE 36, CALCIUM-BINDING.
CC	-/- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.	FT	DOMAIN	2166	2205	EGF-LIKE 37, CALCIUM-BINDING.
CC	-----	FT	DOMAIN	2206	2246	EGF-LIKE 38, CALCIUM-BINDING.
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration	FT	DOMAIN	2247	2290	EGF-LIKE 39, CALCIUM-BINDING.
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	FT	DOMAIN	2291	2332	EGF-LIKE 40, CALCIUM-BINDING.
CC	the European Bioinformatics Institute. There are no restrictions on its	FT	REPEAT	2333	2401	TGFBP 7.
CC	use by non-profit institutions as long as its content is in no way	FT	DOMAIN	2402	2443	EGF-LIKE 41, CALCIUM-BINDING.
CC	modified and this statement is not removed. Usage by and for commercial	FT	DOMAIN	2444	2484	EGF-LIKE 42, CALCIUM-BINDING.
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	FT	DOMAIN	2485	2523	EGF-LIKE 43, CALCIUM-BINDING.
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	FT	DOMAIN	2524	2566	EGF-LIKE 44, CALCIUM-BINDING.
CC	-----	FT	DOMAIN	2567	2606	EGF-LIKE 45, CALCIUM-BINDING.
DR	EMBL; AF073800; AAD50328.1; .	FT	DOMAIN	2607	2647	EGF-LIKE 46, CALCIUM-BINDING.
DR	HSSP; P35555; IAPJ.	FT	DOMAIN	2648	2687	EGF-LIKE 47, CALCIUM-BINDING.
DR	InterPro; IPR000152; Asx_hydroxyl.	FT	DISULFID	85	94	BY SIMILARITY.
DR	InterPro; IPR000561; EGF-like.	FT	DISULFID	89	100	BY SIMILARITY.
DR	InterPro; IPR001881; EGF_Ca.	FT	DISULFID	102	111	BY SIMILARITY.
DR	InterPro; IPR001438; EGF-II.	FT	DISULFID	119	129	BY SIMILARITY.
DR	InterPro; IPR002212; Fibril-assoc.	FT	DISULFID	123	134	BY SIMILARITY.
DR	Pfam; PF00008; EGF; 45.	FT	DISULFID	136	145	BY SIMILARITY.
DR	Pfam; PF00683; TB; 9.	FT	DISULFID	150	160	BY SIMILARITY.
DR	PRINTS; PR00010; EGFBLD.	FT	DISULFID	154	166	BY SIMILARITY.
DR	SMART; SM00179; EGF_Ca; 40.	FT	DISULFID	168	177	BY SIMILARITY.
DR	SMART; SM00001; EGF_like; 6.	FT	DISULFID	250	262	BY SIMILARITY.
DR	PROSITE; PS00010; ASX_HYDROXYL; 41.	FT	DISULFID	257	271	BY SIMILARITY.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.	FT	DISULFID	273	286	BY SIMILARITY.
DR	PROSITE; PS01186; EGF_2; 36.	FT	DISULFID	292	304	BY SIMILARITY.
DR	PROSITE; PS01187; EGF_Ca; 41.	FT	DISULFID	299	313	BY SIMILARITY.
KW	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;	FT	DISULFID	315	328	BY SIMILARITY.
KW	Repeat; Signal; Multigene family.	FT	DISULFID	453	465	BY SIMILARITY.
FT	SIGNAL 1 27	FT	DISULFID	460	474	BY SIMILARITY.
FT	CHAIN 28 2871	FT	DISULFID	476	488	BY SIMILARITY.
FT	DOMAIN 81 112	FT	DISULFID	494	504	BY SIMILARITY.
FT	DOMAIN 115 146	FT	DISULFID	499	513	BY SIMILARITY.
FT	DOMAIN 147 178	FT	DISULFID	515	528	BY SIMILARITY.
FT	DOMAIN 246 287	FT	DISULFID	534	546	BY SIMILARITY.
FT	DOMAIN 288 329	FT	DISULFID	541	555	BY SIMILARITY.
FT	REPEAT 330 390	FT	DISULFID	557	570	BY SIMILARITY.
FT	DOMAIN 392 446	FT	DISULFID	576	587	BY SIMILARITY.
FT	DOMAIN 449 489	FT	DISULFID	582	596	BY SIMILARITY.
FT	DOMAIN 490 529	FT	DISULFID	598	611	BY SIMILARITY.
FT	DOMAIN 530 571	FT	DISULFID	617	628	BY SIMILARITY.
FT	DOMAIN 572 612	FT	DISULFID	623	637	BY SIMILARITY.
FT	DOMAIN 613 653	FT	DISULFID	639	652	BY SIMILARITY.
FT	REPEAT 654 722	FT	DISULFID	727	739	BY SIMILARITY.
FT	DOMAIN 723 764	FT	DISULFID	734	748	BY SIMILARITY.
FT	DOMAIN 765 806	FT	DISULFID	750	763	BY SIMILARITY.
FT	DOMAIN 807 846	FT	DISULFID	769	781	BY SIMILARITY.
FT	DOMAIN 910 951	FT	DISULFID	776	790	BY SIMILARITY.
FT	REPEAT 952 1027	FT	DISULFID	792	805	BY SIMILARITY.









QY 302 -----QPSYSCHMETGYRLAADQHRCEVDVDCILEPSPCP-QRCVNTQGGFECHCYP 353  
DB 1294 DGGQCTNIPGEGYRLCYDGFMSDMKTCIDVNECDLNSNCFMGECENTKGSFICHQCL 1353  
QY 354 NYDLVDGE-CVEVDPC--FRANCEVQCQPLN-QTSVLCVCAEGFAPIPH-----EPHR 403  
DB 1354 GYSVKKGTTGCTD-VDECEIGAHCNDMHSCLNIPGSKSCREG--IGNGIKCIDLDE 1410  
QY 404 COMFCNQATCAPDNPQAS--CECPGTYLDDGFTCTDIDE-----CENG----- 448  
DB 1411 CSNGTHQCSINACQ-VNTPGYSRCACSEGP-TGDGFTCSVDDECAENINLNCNGCQCLNVP 1468  
QY 449 -----GF-----CSGVCHNLPGTFECICGPDPSALARHIG- 477  
DB 1469 GAYRCECEMGTFPASDSRSCQDIDECFSQNICVSGTCNNLPGMFHCICDDGYELDRGTGN 1528  
QY 478 -TDGD 481  
DB 1529 CTDDID 1533  
RESULT 12  
FBN1\_HUMAN  
ID FBN1\_HUMAN STANDARD; PRT; 2871 AA.  
AC P35555;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrillin 1 precursor.  
GN FBN1 OR FBN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=93372860; PubMed=8364578;  
RA Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,  
Pangillan T., Bonadio J.;  
RT "Genomic organization of the sequence coding for fibrillin, the  
defective gene product in Marfan syndrome.";  
RL Hum. Mol. Genet. 2:961-968(1993).  
RN [2]  
RP SEQUENCE OF 1-932 FROM N.A.  
RC TISSUE=Placenta, and Fibroblast;  
RX MEDLINE=94010947; PubMed=7691719;  
RA Corson G.M., Chabberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;  
RT "Fibrillin binds calcium and is coded by cDNAs that reveal a  
multidomain structure and alternatively spliced exons at the 5'  
end.";  
RL Genomics 17:476-484(1993).  
RN [3]  
RP SEQUENCE OF 899-2871 FROM N.A.  
RX MEDLINE=91304566; PubMed=1852207;  
RA Maslen C.L., Corson G.M., Maddox B.K., Glanville R.W., Sakai L.Y.;  
RT "Partial sequence of a candidate gene for the Marfan syndrome.";  
RL Nature 352:334-337(1991).  
RN [4]  
RP SEQUENCE OF 813-1313 FROM N.A.  
RX MEDLINE=91304567; PubMed=1852206;  
RA Lee B., Godfrey M., Vitale E., Hori H., Mattel M.-G., Sarfarazi M.,  
Tsiouras P., Ramirez F., Hollister D.W.;  
RT "Linkage of Marfan syndrome and a phenotypically related disorder to  
two different fibrillin genes.";  
RL Nature 352:330-334(1991).  
RN [5]  
RP CHARACTERIZATION.  
RX MEDLINE=91317849; PubMed=1860873;  
RA Sakai L.Y., Keene D.R., Glanville R.W., Bachinger H.P.;  
RT "Purification and partial characterization of fibrillin, a cysteine-  
rich structural component of connective tissue microfibrils.";

RL J. Biol. Chem. 266:14763-14770(1991).  
RN [6]  
RP STRUCTURE BY NMR OF 2054-2125.  
RX MEDLINE=98031893; PubMed=9362480;  
RA Yuan X., Downing A.K., Knott V., Handford P.A.;  
RT "Solution structure of the transforming growth factor beta-binding  
protein-like module, a domain associated with matrix fibrils.";  
RL EMBO J. 16:6659-6666(1997).  
RN [7]  
RP STRUCTURE BY NMR OF 2124-2205.  
RX MEDLINE=96144829; PubMed=8568869;  
RA Knott V., Downing A.K., Cardy C.M., Handford P.A.;  
RT "Calcium binding properties of an epidermal growth factor-like domain  
from human fibrillin-1.";  
RL J. Mol. Biol. 255:22-27(1996).  
RN [8]  
RP STRUCTURE BY NMR OF 2124-2205.  
RX MEDLINE=96222301; PubMed=8653794;  
RA Downing A.K., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,  
Handford P.A.;  
RT "Solution structure of a pair of calcium-binding epidermal growth  
factor-like domains: implications for the Marfan syndrome and other  
genetic disorders.";  
RL Cell 85:597-605(1996).  
RN [9]  
RP REVIEW ON MFS VARIANTS.  
RX MEDLINE=96174615; PubMed=8594563;  
RA Colod G., Beroud C., Soussi T., Junien C., Boileau C.;  
RT "Software and database for the analysis of mutations in the human  
FBN1 gene.";  
RL Nucleic Acids Res. 24:137-141(1996).  
RN [10]  
RP REVIEW ON MFS VARIANTS.  
RX MEDLINE=97169383; PubMed=9016526;  
RA Colod-Beroud G., Beroud C., Ades L., Black C., Boxer M., Brock D.J.,  
Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,  
Richards R.J., Wang W., Junien C., Boileau C.;  
RT "Marfan Database (second edition): software and database for the  
analysis of mutations in the human FBN1 gene.";  
RL Nucleic Acids Res. 25:147-150(1997).  
RN [11]  
RP REVIEW ON VARIANTS.  
RX MEDLINE=98062175; PubMed=9401003;  
RA Hayward C., Brock D.J.H.;  
RT "Fibrillin-1 mutations in Marfan syndrome and other type-1  
fibrillinopathies.";  
RL Hum. Mutat. 10:415-423(1997).  
RN [12]  
RP VARIANT MFS PRO-1137.  
RX MEDLINE=91304569; PubMed=1852208;  
RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,  
Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,  
Curristin S.M., Stetten G., Meyers D.A., Francomano C.A.;  
RT "Marfan syndrome caused by a recurrent de novo missense mutation in  
the fibrillin gene.";  
RL Nature 352:337-339(1991).  
RN [13]  
RP VARIANTS MFS SER-1249; ARG-1663; SER-2221 AND SER-2307.  
RX MEDLINE=93250834; PubMed=1301946;  
RA Dietz H.C., Saraiva J.M., Pyeritz R.E., Cutting G.R., Francomano C.A.;  
RT "Clustering of fibrillin (FBN1) missense mutations in Marfan syndrome  
patients at cysteine residues in EGF-like domains.";  
RL Hum. Mutat. 1:366-374(1992).  
RN [14]  
RP VARIANT MFS SER-2307.  
RX MEDLINE=92235290; PubMed=1569206;  
RA Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J. Jr.,  
Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;  
RT "Marfan phenotype variability in a family segregating a missense  
mutation in the epidermal growth factor-like motif of the fibrillin  
gene.";  
RL J. Clin. Invest. 89:1674-1680(1992).  
RN [15]



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	DR	EMBL; L29454; AAA56840.1; -	FT	DOMAIN	2402	2443	EGF-LIKE 41, CALCIUM-BINDING.
CC	DR	EMBL; U22493; AAA64217.1; -	FT	DOMAIN	2444	2484	EGF-LIKE 42, CALCIUM-BINDING.
CC	DR	HSSP; P35555; IAPJ.	FT	DOMAIN	2485	2523	EGF-LIKE 43, CALCIUM-BINDING.
CC	DR	MCD; MGI:95489; Fbn1.	FT	DOMAIN	2524	2566	EGF-LIKE 44, CALCIUM-BINDING.
DR	DR	InterPro; IPR000152; Asx_hydroxyl.	FT	DOMAIN	2567	2606	EGF-LIKE 45, CALCIUM-BINDING.
DR	DR	InterPro; IPR001881; EGF_Ca.	FT	DOMAIN	2607	2647	EGF-LIKE 46, CALCIUM-BINDING.
DR	DR	InterPro; IPR002212; Fibril-assoc.	FT	DOMAIN	2648	2687	EGF-LIKE 47, CALCIUM-BINDING.
DR	DR	Pfam; PF00008; EGF; 46.	FT	DOMAIN	85	94	BY SIMILARITY.
DR	DR	Pfam; PF00683; TB; 9.	FT	DISULFID	89	100	BY SIMILARITY.
DR	DR	SMART; SM00179; EGF_CA; 42.	FT	DISULFID	102	111	BY SIMILARITY.
DR	DR	SMART; SM00001; EGF_like; 4.	FT	DISULFID	119	129	BY SIMILARITY.
DR	DR	PROSITE; PS00010; ASX_HYDROXYL; 43.	FT	DISULFID	123	134	BY SIMILARITY.
DR	DR	PROSITE; PS00022; EGF_1; 2.	FT	DISULFID	136	145	BY SIMILARITY.
DR	DR	PROSITE; PS01186; EGF_1; 2.	FT	DISULFID	150	160	BY SIMILARITY.
DR	DR	PROSITE; PS01187; EGF_CA; 43.	FT	DISULFID	168	177	BY SIMILARITY.
DR	DR	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain; Repeat; Signal; Multigene family.	FT	DISULFID	177	186	BY SIMILARITY.
FT	FT	SIGNAL	FT	DISULFID	250	262	BY SIMILARITY.
FT	FT	POTENTIAL.	FT	DISULFID	271	286	BY SIMILARITY.
FT	FT	CHAIN	FT	DISULFID	292	304	BY SIMILARITY.
FT	FT	28	FT	DISULFID	299	313	BY SIMILARITY.
FT	FT	81	FT	DISULFID	315	328	BY SIMILARITY.
FT	FT	115	FT	DISULFID	328	345	BY SIMILARITY.
FT	FT	146	FT	DISULFID	345	365	BY SIMILARITY.
FT	FT	178	FT	DISULFID	453	474	BY SIMILARITY.
FT	FT	246	FT	DISULFID	474	488	BY SIMILARITY.
FT	FT	288	FT	DISULFID	494	504	BY SIMILARITY.
FT	FT	330	FT	DISULFID	499	513	BY SIMILARITY.
FT	FT	401	FT	DISULFID	515	528	BY SIMILARITY.
FT	FT	446	FT	DISULFID	534	546	BY SIMILARITY.
FT	FT	489	FT	DISULFID	541	555	BY SIMILARITY.
FT	FT	490	FT	DISULFID	557	570	BY SIMILARITY.
FT	FT	529	FT	DISULFID	570	587	BY SIMILARITY.
FT	FT	530	FT	DISULFID	582	596	BY SIMILARITY.
FT	FT	571	FT	DISULFID	598	611	BY SIMILARITY.
FT	FT	612	FT	DISULFID	617	628	BY SIMILARITY.
FT	FT	613	FT	DISULFID	623	637	BY SIMILARITY.
FT	FT	653	FT	DISULFID	639	652	BY SIMILARITY.
FT	FT	721	FT	DISULFID	727	739	BY SIMILARITY.
FT	FT	723	FT	DISULFID	734	748	BY SIMILARITY.
FT	FT	765	FT	DISULFID	750	763	BY SIMILARITY.
FT	FT	807	FT	DISULFID	769	781	BY SIMILARITY.
FT	FT	846	FT	DISULFID	776	790	BY SIMILARITY.
FT	FT	910	FT	DISULFID	792	805	BY SIMILARITY.
FT	FT	951	FT	DISULFID	811	821	BY SIMILARITY.
FT	FT	1018	FT	DISULFID	816	830	BY SIMILARITY.
FT	FT	1028	FT	DISULFID	832	845	BY SIMILARITY.
FT	FT	1069	FT	DISULFID	845	858	BY SIMILARITY.
FT	FT	1112	FT	DISULFID	858	871	BY SIMILARITY.
FT	FT	1154	FT	DISULFID	871	884	BY SIMILARITY.
FT	FT	1197	FT	DISULFID	884	897	BY SIMILARITY.
FT	FT	1237	FT	DISULFID	897	910	BY SIMILARITY.
FT	FT	1279	FT	DISULFID	910	923	BY SIMILARITY.
FT	FT	1321	FT	DISULFID	923	936	BY SIMILARITY.
FT	FT	1362	FT	DISULFID	936	949	BY SIMILARITY.
FT	FT	1403	FT	DISULFID	949	962	BY SIMILARITY.
FT	FT	1445	FT	DISULFID	962	975	BY SIMILARITY.
FT	FT	1486	FT	DISULFID	975	988	BY SIMILARITY.
FT	FT	1527	FT	DISULFID	988	1001	BY SIMILARITY.
FT	FT	1599	FT	DISULFID	1001	1014	BY SIMILARITY.
FT	FT	1647	FT	DISULFID	1014	1027	BY SIMILARITY.
FT	FT	1688	FT	DISULFID	1027	1040	BY SIMILARITY.
FT	FT	1758	FT	DISULFID	1040	1053	BY SIMILARITY.
FT	FT	1807	FT	DISULFID	1053	1066	BY SIMILARITY.
FT	FT	1848	FT	DISULFID	1066	1079	BY SIMILARITY.
FT	FT	1890	FT	DISULFID	1079	1092	BY SIMILARITY.
FT	FT	1929	FT	DISULFID	1092	1105	BY SIMILARITY.
FT	FT	1972	FT	DISULFID	1105	1118	BY SIMILARITY.
FT	FT	2012	FT	DISULFID	1118	1131	BY SIMILARITY.
FT	FT	2054	FT	DISULFID	1131	1144	BY SIMILARITY.
FT	FT	2121	FT	DISULFID	1144	1157	BY SIMILARITY.
FT	FT	2165	FT	DISULFID	1157	1170	BY SIMILARITY.
FT	FT	2205	FT	DISULFID	1170	1183	BY SIMILARITY.
FT	FT	2246	FT	DISULFID	1183	1196	BY SIMILARITY.
FT	FT	2290	FT	DISULFID	1196	1209	BY SIMILARITY.
FT	FT	2332	FT	DISULFID	1209	1222	BY SIMILARITY.
FT	FT	2400	FT	DISULFID	1222	1235	BY SIMILARITY.
FT	FT		FT	DISULFID	1235	1248	BY SIMILARITY.
FT	FT		FT	DISULFID	1248	1261	BY SIMILARITY.
FT	FT		FT	DISULFID	1261	1274	BY SIMILARITY.
FT	FT		FT	DISULFID	1274	1287	BY SIMILARITY.
FT	FT		FT	DISULFID	1287	1300	BY SIMILARITY.
FT	FT		FT	DISULFID	1300	1313	BY SIMILARITY.
FT	FT		FT	DISULFID	1313	1326	BY SIMILARITY.
FT	FT		FT	DISULFID	1326	1339	BY SIMILARITY.
FT	FT		FT	DISULFID	1339	1352	BY SIMILARITY.
FT	FT		FT	DISULFID	1352	1365	BY SIMILARITY.

FT DISULFID 1367 1380 BY SIMILARITY.  
FT DISULFID 1374 1389 BY SIMILARITY.  
FT DISULFID 1391 1402 BY SIMILARITY.  
FT DISULFID 1408 1420 BY SIMILARITY.  
FT DISULFID 1415 1429 BY SIMILARITY.  
FT DISULFID 1431 1444 BY SIMILARITY.  
FT DISULFID 1450 1461 BY SIMILARITY.  
FT DISULFID 1456 1470 BY SIMILARITY.  
FT DISULFID 1472 1485 BY SIMILARITY.  
FT DISULFID 1491 1502 BY SIMILARITY.  
FT DISULFID 1497 1511 BY SIMILARITY.  
FT DISULFID 1513 1526 BY SIMILARITY.  
FT DISULFID 1610 1622 BY SIMILARITY.  
FT DISULFID 1617 1631 BY SIMILARITY.  
FT DISULFID 1633 1646 BY SIMILARITY.  
FT DISULFID 1652 1663 BY SIMILARITY.  
FT DISULFID 1658 1672 BY SIMILARITY.  
FT DISULFID 1674 1687 BY SIMILARITY.  
  
Query Match 12.4%; Score 362; DB 1; Length 2871;  
Best Local Similarity 29.4%; Pred. No. 2.5e-16;  
Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;  
  
QY 212 SAAVAPLGL-----QLMCTAPPAGVGHVAREAPGAWDCSVENGCGEHCACNAIPGAPRC 265  
DB 1162 SANLCPHRCVNLGKYOCACNPGVHPHDLFCVVIDECSIMNGCCTFTCTNSDGSYEC 1221  
  
QY 266 QCPAGALQADGRSCTASATOSCNLCEHFCVNP-----DQPGSYSCMCETGYRL 316  
DB 1222 SCQPFALMPDQSRCT-----DIDQ--CEDNPICDGGCTNIPGEYRCLCYDGPMA 1271  
  
QY 317 AADQHRCEVDVDCILEPFC--PQRCVNTQGFECYHCYNYDLVGE--CVPVDPDPC--FR 371  
DB 1272 SEDMKTCVDVNECDLNPICLSGTCTNKSGFICHDGMYSGKGTGCTD--INECEIGA 1330  
  
QY 372 ANCEYOCPLOT--SYLCVCAEGFA-----PIPEPHRCQMFQNCQACPADCPN 420  
DB 1331 HNCGRHACTNTAGSFCKSPGNGIDGKICTDLDECSNGTHMSQH-----ADC-KN 1382  
  
QY 421 TQAS--CCCEGYLDDGFTCTIDDEC--EN-----GGF----- 450  
DB 1383 TMSGYRCLCKDGY--TGDFCTCTDLDECSNENLNGGQCLNAPGGYRCECDMGVPSADG 1441  
  
QY 451 -----CS-----GVCNHLPTFCICGPDPSALARHIG-----TDCDGGK 484  
DB 1442 KACEDIDECSLPNICVFTCHNLPLGFRCECEIGYELDRSGGNCCTVNECLDPTTCTISGN 1501  
  
QY 485 VDGDSGSGEPSPPTGGS-----TLPPPAVGLV 513  
DB 1502 CVN-----TPGSYTCDCSPDELPNTRVGCV 1527  
  
RESULT 14  
FBL2\_MOUSE  
ID FBL2\_MOUSE STANDARD; PRF; 1221 AA.  
AC P37889; Q9WU12;  
DT 01-Oct-1994 (Rel. 30, Created)  
DT 01-Oct-1994 (Rel. 30, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibulin-2 precursor.  
GN FBLN2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.  
RC TISSUE=Fibroblast;  
RX MEDLINE=94064787; PubMed=8245130;  
RA Pan T.-C., Sakaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;  
RT "Structure and expression of fibulin-2, a novel extracellular matrix  
RT protein with multiple EGF-like repeats and consensus motifs for  
RT calcium binding.";

J. Cell Biol. 123:1269-1277(1993).  
[2]  
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
MEDLINE=9937686; PubMed=10406956;  
Graessel S., Sicot F.-X., Gotta S., Chu M.-L.;  
"Mouse fibulin-2 gene. Complete exon-intron organization and promoter  
characterization.";  
Eur. J. Biochem. 263:471-477(1999).  
CC -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS  
CALCIUM DEPENDENT.  
CC -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.  
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.  
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms; 1 (shown here) and  
2/EGF3-less; are produced by alternative splicing.  
CC -1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER  
CONNECTIVE TISSUES.  
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X75285; CAA53040.1; -  
EMBL; AF135253; AAD34456.1; -  
EMBL; AF135239; AAD34456.1; JOINED.  
EMBL; AF135240; AAD34456.1; JOINED.  
EMBL; AF135241; AAD34456.1; JOINED.  
EMBL; AF135242; AAD34456.1; JOINED.  
EMBL; AF135243; AAD34456.1; JOINED.  
EMBL; AF135244; AAD34456.1; JOINED.  
EMBL; AF135245; AAD34456.1; JOINED.  
EMBL; AF135246; AAD34456.1; JOINED.  
EMBL; AF135247; AAD34456.1; JOINED.  
EMBL; AF135248; AAD34456.1; JOINED.  
EMBL; AF135249; AAD34456.1; JOINED.  
EMBL; AF135250; AAD34456.1; JOINED.  
EMBL; AF135251; AAD34456.1; JOINED.  
EMBL; AF135252; AAD34456.1; JOINED.  
PIR; A49457; A49457.  
HSP; P00736; IAPQ.  
MGD; MGI:95488; Fbln2.  
InterPro; IPR000020; Anaphylatoxin.  
InterPro; IPR000152; Asx\_hydroxyl.  
InterPro; IPR000561; EGF-like.  
InterPro; IPR001881; EGF\_Ca.  
Pfam; PF00008; EGF; 6.  
Pfam; PF01821; ANATO; 2.  
SMART; SM00104; ANATO; 3.  
SMART; SM00179; EGF\_CA; 9.  
SMART; SM00001; EGF\_like; 2.  
PROSITE; PS00010; ASX\_HYDROXYL; 5.  
PROSITE; PS01177; ANAPHYLATOXIN\_1; 3.  
PROSITE; PS01178; ANAPHYLATOXIN\_2; 3.  
PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
PROSITE; PS01186; EGF\_2; 5.  
PROSITE; PS01187; EGF\_CA; 10.  
Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;  
Calcium-binding; Alternative splicing; Repeat.  
FT SIGNAL 1 26  
FT CHAIN 27 1221 FIBULIN-2.  
FT DOMAIN 27 434 N.  
FT DOMAIN 27 176 SUBDOMAIN NA (CYS-RICH).  
FT DOMAIN 177 434 SUBDOMAIN NB (CYS-FREE).  
FT DOMAIN 435 477 ANAPHYLATOXIN-LIKE 1.  
FT DOMAIN 478 510 ANAPHYLATOXIN-LIKE 2.  
FT DOMAIN 511 543 ANAPHYLATOXIN-LIKE 3.  
FT DOMAIN 594 635 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 669 708 EGF-LIKE 2.

```

FT DOMAIN 709 755 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 756 800 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 801 846 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 847 894 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 895 937 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 938 979 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 980 1018 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1019 1061 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1062 1106 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1111 1221 DOMAIN III.
FT SITE 421 423 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 435 462 BY SIMILARITY.
FT DISULFID 436 469 BY SIMILARITY.
FT DISULFID 449 470 BY SIMILARITY.
FT DISULFID 479 508 BY SIMILARITY.
FT DISULFID 492 509 BY SIMILARITY.
FT DISULFID 511 535 BY SIMILARITY.
FT DISULFID 512 543 BY SIMILARITY.
FT DISULFID 525 542 BY SIMILARITY.
FT DISULFID 598 610 BY SIMILARITY.
FT DISULFID 606 619 BY SIMILARITY.
FT DISULFID 621 634 BY SIMILARITY.
FT DISULFID 673 683 BY SIMILARITY.
FT DISULFID 679 692 BY SIMILARITY.
FT DISULFID 694 707 BY SIMILARITY.
FT DISULFID 713 726 BY SIMILARITY.
FT DISULFID 720 735 BY SIMILARITY.
FT DISULFID 742 754 BY SIMILARITY.
FT DISULFID 805 818 BY SIMILARITY.
FT DISULFID 812 827 BY SIMILARITY.
FT DISULFID 833 845 BY SIMILARITY.
FT DISULFID 899 912 BY SIMILARITY.
FT DISULFID 906 921 BY SIMILARITY.
FT DISULFID 923 936 BY SIMILARITY.
FT DISULFID 942 954 BY SIMILARITY.
FT DISULFID 950 963 BY SIMILARITY.
FT DISULFID 965 978 BY SIMILARITY.
FT DISULFID 984 993 BY SIMILARITY.
FT DISULFID 989 1002 BY SIMILARITY.
FT DISULFID 1004 1017 BY SIMILARITY.
FT DISULFID 1023 1035 BY SIMILARITY.
FT DISULFID 1031 1044 BY SIMILARITY.
FT DISULFID 1046 1060 BY SIMILARITY.
FT DISULFID 1066 1079 BY SIMILARITY.
FT DISULFID 1073 1088 BY SIMILARITY.
FT DISULFID 1093 1105 BY SIMILARITY.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 709 755 MISSING (IN ISOFORM 2).
FT CONFLICT 140 159 HSGRYAAGHTVHLSSCRAC -> TVAVSICWYRPLPLIP
FT CONFLICT 348 348 S -> L (IN REF. 2).
FT CONFLICT 507 507 Q -> QO (IN REF. 2).
FT CONFLICT 1102 1102 Q -> E (IN REF. 2).
SQ SEQUENCE 1221 AA; 131818 MW; 87DB2A10A8FDC45F CRC64;

Query Match
Best Local Similarity 29.5%; Pred. No. 2.4e-16;
Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;

Qy 154 EQQCEVKAD-GFTCEHFPATC-----RPLAV-----EPGAAARVS-----IT 192
Db 521 EGQSCSNPLGYPCN-HVMSLSCGEPLTVPEVRPPPEAAPPVSEMAASREALS 579
Qy 193 YGT---PFAARGADFO---ALP-----VGSAAVAPLGLQLM-----CTAPPG 229
Db 580 LGTEALPNSLPQDDQDECLMLPGELCHLCINTVSYRCACPFPGFELQDGRTRCDRPG 639
Qy 230 AVQGHWARE-APGAWDCSV-----ENGCEHACNAIPGAPRQCQCPAGAA 272
Db 640 APQLDTARESAPRESAQVSPNTIPLVPQPNCTCKDNGPCRCVGVGDTAMCSFPGYA 699

```







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:22:32 ; Search time 31.5 Seconds  
(without alignments)  
3375.248 Million cell updates/sec

Title: US-09-509-994-2

Perfect score: 2916  
Sequence: 1 MLGVLVIGALALAGLGFAP.....PSPTPGSTLTPPAVLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	2658	91.2	468	4 Q9UC32	Q9uc32 homo sapien
2	1844	63.2	577	11 Q35370	Q35370 rattus norv
3	1400.5	48.0	461	11 P97883	P97883 rattus norv
4	584.5	20.0	757	4 Q9HC00	Q9hcu0 homo sapien
5	575	19.7	765	11 Q91ZV1	Q91zv1 mus musculu
6	575	19.7	765	11 Q91V98	Q91v98 mus musculu
7	431	14.8	1664	5 Q9TVQ2	Q9tvg2 caenorhabdi
8	414	14.2	1574	11 Q88281	Q88281 rattus norv
9	386	13.2	1394	5 Q9VS89	Q9vs89 drosophila
10	376	12.9	708	13 P87363	P87363 gallus gall
11	375	12.9	2906	11 Q9WU9H	Q9wuh9 rattus norv
12	369	12.7	1246	4 Q75095	Q75095 homo sapien
13	367	12.6	3857	11 Q88840	Q88840 mus musculu
14	365	12.5	2809	4 Q96JP8	Q96jp8 homo sapien
15	361	12.4	2872	11 Q9WU98	Q9wuh8 rattus norv
16	359	12.3	528	11 Q9CXD8	Q9cxd8 mus musculu

17	358.5	12.3	1174	11 Q99K58	Q99k58 mus musculu
18	357	12.2	937	4 Q96FT5	Q96ft5 homo sapien
19	353.5	12.1	746	4 Q96HB9	Q96hb9 homo sapien
20	353.5	12.1	1256	4 Q9NS15	Q9ns15 homo sapien
21	353.5	12.1	1382	4 Q9H7K2	Q9h7k2 homo sapien
22	352.5	12.1	741	4 Q96R89	Q96r89 homo sapien
23	351.5	12.1	1511	4 Q75412	Q75412 homo sapien
24	351.5	12.1	1587	4 Q00508	Q00508 homo sapien
25	342	11.7	5636	4 Q96RW7	Q96rw7 homo sapien
26	341	11.7	576	4 Q9Y3V7	Q9y3v7 homo sapien
27	340	11.7	1062	11 Q60789	Q60789 mus musculu
28	338.5	11.6	1289	5 Q8SSS3	Q8sss3 dictyosteli
29	336	11.5	2673	4 Q96SC3	Q96sc3 homo sapien
30	333.5	11.4	961	11 Q9BQC6	Q9eqc6 mus musculu
31	332.5	11.4	1764	11 Q35806	Q35806 rattus norv
32	332	11.4	999	4 Q9NQ36	Q9nq36 homo sapien
33	331.5	11.4	1713	11 Q88349	Q88349 mus musculu
34	331	11.4	589	5 Q9T2S1	Q9t2s1 caenorhabdi
35	329.5	11.3	956	11 Q8R542	Q8r542 mus musculu
36	329	11.3	2189	5 Q9BI05	Q9bi05 eimeria ten
37	328.5	11.3	956	11 Q99K64	Q99k64 mus musculu
38	325	11.1	1095	11 Q60784	Q60784 mus musculu
39	324.5	11.1	937	5 Q9BLJ1	Q9blj1 ciona intes
40	323	11.1	1253	11 Q61810	Q61810 mus musculu
41	322.5	11.1	1394	4 Q8TD95	Q8td95 homo sapien
42	319.5	11.0	1833	11 Q08999	Q08999 mus musculu
43	315.5	10.8	443	4 Q96TF5	Q96tf5 homo sapien
44	314.5	10.8	3507	5 Q23587	Q23587 caenorhabdi
45	314	10.8	997	11 Q9JJS0	Q9jjjs0 mus musculu

ALIGNMENTS

RESULT 1

Q9UC32	PRELIMINARY;	PRT;	468 AA.
ID	Q9UC32		
AC	Q9UC32;		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)		
DE	THROMBOMODULIN.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=93293792; PubMed=8390446;		
RA	Yamanoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;		
RT	"Urinary thrombomodulin, its isolation and characterization.";		
RL	J. Biochem. 113:433-440(1993).		
DR	HSP; P07204; IZAQ.		
DR	InterPro; IPR000152; Asx_hydroxyl.		
DR	InterPro; IPR000561; EGF-like.		
DR	InterPro; IPR001881; EGF_Ca.		
DR	InterPro; IPR001304; Lectin_C.		
DR	InterPro; IPR001491; Thrombomoduln.		
DR	Pfam; PF00008; EGF_4.		
DR	Pfam; PF00059; Lectin_C; 1.		
DR	PRINTS; PR00907; THROMBOMODULN.		
DR	SMART; SM00034; CLECT; 1.		
DR	SMART; SM0181; EGF; 6.		
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.		
DR	PROSITE; PS00041; C-TYPE_LECTIN_2; 1.		
DR	PROSITE; PS01186; EGF_2; 2.		
DR	PROSITE; PS01187; EGF_CA; 1.		
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.		
SQ	SEQUENCE 468 AA; 49444 MW; 4BFEBE98EFB86A40 CRC64;		

Query Match 91.2%; Score 2658; DB 4; Length 468;  
Best Local Similarity 99.8%; Pred. NO. 8.4e-205;  
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 APAEPQGGSCVCEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 78  
 Db 1 APAEPQGGSCVCEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
 QY 79 GVGRRRLWIGLQPLPGCGDPKRLGRLGFGQWGTGDNNTSYSRWARLDLNGAPLCGPLCV 138  
 Db 61 GVGRRRLWIGLQPLPGCGDPKRLGRLGFGQWGTGDNNTSYSRWARLDLNGAPLCGPLCV 120  
 QY 139 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFPFCRPLAVEPFGAAAAVSYTGTPFA 198  
 Db 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFPFCRPLAVEPFGAAAAVSYTGTPFA 180  
 QY 199 ARGADFQALPVGSSAAVAPLGLQMLCTAPGAVQGHAREAPGAWDCSVENGCEHACNA 258  
 Db 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPGAVQGHAREAPGAWDCSVENGCEHACNA 240  
 QY 259 IPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCEGYRLAA 318  
 Db 241 IPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCEGYRLAA 300  
 QY 319 DQHRCEVDVDCILEPSPCQRCVNTQGGFCHCYPNVDLVGCEVPEVDPFCFRANCEYQC 378  
 Db 301 DQHRCEVDVDCILEPSPCQRCVNTQGGFCHCYPNVDLVGCEVPEVDPFCFRANCEYQC 360  
 QY 379 QPLNQTSLVLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECEGYILDDGFI 438  
 Db 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECEGYILDDGFI 420  
 QY 439 CTDIDECENGFCGVCVCHNLPCTGTECICGPDSSALAHIGTDCDSGKYD 486  
 Db 421 CTDIDECENGFCGVCVCHNLPCTGTECICGPDSSALAHIGTDCDSGKYD 468  
 RESULT 2  
 ID Q35370 PRELIMINARY; PRT; 577 AA.  
 AC Q35370;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE THROMBOMODULIN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;  
 RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-136 FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF022743; AAB80760.1; -;  
 DR EMBL; AF022742; AAB80923.1; -;  
 DR HSSP; P07204; 1FGD.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001304; Lectin\_C.  
 DR InterPro; IPR001491; Thrombomodulin.  
 DR Pfam; PF00008; EGF; 5.  
 DR Pfam; PF00059; Lectin\_C; 1.  
 DR PRINTS; PR00907; THROMBOMODULN.  
 DR SMART; SM00034; CLECT; 1.  
 DR SMART; SM00179; EGF\_Ca; 2.  
 DR SMART; SM00001; EGF\_Like; 3.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 2.  
 DR PROSITE; PS00041; C-TYPE\_LECTIN\_2; 1.  
 DR PROSITE; PS01186; EGF\_2; 3.

DR PROSITE; PS01187; EGF\_CA; 2.  
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 SQ SEQUENCE 577 AA; 61844 MW; 0BE764CBF1B555F CRC64;  
 Query Match .63.2%; Score 1844; DB 11; Length 577;  
 Best Local Similarity 64.0%; Pred. No. 1.6e-139;  
 Matches 333; Conservative 49; Mismatches 133; Indels 6; Gaps 5;  
 QY 1 MLGVLVGLALAGLGFAPAPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 Db 1 MLGVLVGLAPAGLGLSALAKLPKQSCVGNCFALFQDPVTELDASQACORLQGHLM 60  
 QY 61 TVRSSVAADVLSLLNGDGVGRRRLWIGLQPLPGCGDPKRLGRLGFGQWGTGDNNTSYS 120  
 Db 61 TVRSSVAADVLSLLVS-DSSMDSRP-WIGLQPLPGCGDPVHLGRLGFGQWGTGDNNTSYS 118  
 QY 121 RWARLDLNGAPLCGPLCVAAVSAEATVPSEPIWEEOQCEVKADGFLCEHFFPFCRPLAV 180  
 Db 119 RWARNDSQSPPLCGPLCVTVSTATAAFCBPANEEKPCENETKGLCFEYFAAFCRPLAV 178  
 QY 181 EP-GAAAAVSYTGTPFAAGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHAREA 239  
 Db 179 NTRDPEGAHISSTYNTPLGVSGADFQTLPGSSATVAPFGLVLCRALPGTSEGHWTREV 238  
 QY 240 PGAWDCSVENGCEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 299  
 Db 239 TGAWNCSVENGCEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 298  
 QY 300 PQPGSYSCMCEGYRLAADHRCEDVDVDCILEPSPCQRCVNTQGGFCHCYPNVDLV 359  
 Db 299 SDVPGSYSCMCEGYRLAADHRCEDVDVDCILEPSPCQRCVNTQGGFCHCYPNVDLV 358  
 QY 360 GECVEVDPCFRANCEYQCPLNQTSLVLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDP 419  
 Db 359 GECVEQLDPCFRSKCEYQCPLNQTSLVLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDP 418  
 QY 420 NTQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPCTGTECICGPDSSALAHIGT 479  
 Db 419 NSPFCQCEGYILDDGFICTDIDECENGFCGVCVCHNLPCTGTECICGPDSSALAHIGT 478  
 QY 480 CDGSGKV--DGGDGGSGGEGPPPS-PTGSEITLTPPAVLGVHSG 516  
 Db 479 CDPIPVLESDGSGGEGHPSSNPVTVSVVPSARPMSHG 518  
 RESULT 3  
 ID P97883 PRELIMINARY; PRT; 461 AA.  
 AC P97883;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Thrombomodulin (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN CAPILLARY;  
 RA Wang L., Tran N.D., Schreiber S.S., Zlokovic B.V.;  
 RL "Nucleotide sequence of rat thrombomodulin."  
 DR EMBL; U90121; AAB49723.1; -;  
 DR HSSP; P07204; 1FGD.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001491; Thrombomodulin.  
 DR Pfam; PF00008; EGF; 5.  
 DR PRINTS; PR00907; THROMBOMODULN.  
 DR SMART; SM00179; EGF\_CA; 2.  
 DR SMART; SM00001; EGF\_Like; 3.





AC Q9TV02;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Y64G10A.7 protein.  
GN Y64G10A.7  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Mortimore B.J.;  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RN RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology."  
RL Science 282:2012-2018(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Ainscough R.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL117206; CAB60454.1; -  
DR EMBL; AL110498; CAB60454.1; JOINED.  
DR EMBL; AL110498; CAB57911.1; -  
DR EMBL; AL117206; CAB57911.1; JOINED.  
DR HSSP; P00736; IAPQ.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR Pfam; PF00008; EGF; 25.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR SMART; SM00179; EGF\_CA; 4.  
DR SMART; SM00001; EGF\_like; 18.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_22.  
DR PROSITE; PS01186; EGF\_2; 24.  
DR PROSITE; PS01187; EGF\_CA; 3.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
SQ SEQUENCE 1664 AA; 179279 MW; A69F093B4C705932 CRC64;

Query Match 14.8%; Score 431; DB 5; Length 1664;  
Best Local Similarity 29.7%; Pred. No. 6.7e-26;  
Matches 126; Conservative 40; Mismatches 178; Indels 80; Gaps 19;

QY 119 YSRWRLDNLGAPLCPLCVAVSAE-----ATVPSEPIWEEQCE--VKADGFLCEF 169  
DB 56 YLRFAFRSRGCKCLLRVQANCSADLCHNGGTCTVPSEHNDNEQVCECPVGTGAKCQY 115  
QY 170 HFPATCRPLAVEPFAAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPP 229  
DB 116 D-ANEC--MANNGGCEHCYN-TIGTY-----CRWPG 145

QY 230 AVQGHWAREAPAGWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTA----- 282  
DB 146 FELSGDGNCTSDIDECAVSGNGCSDRCVNSPGGFRCDPCLYHADRTGSGGFHENL 205

QY 283 ---SATQSC---NDLCEHFCVNPDPGSGYSCMETGYRLADQHRCEVDVDDCILPEPSC 336  
DB 206 ILIKKVTSCSTDNGGCEHC-ENDSNGEYRCRCRVGFKLSENKRCQPVDPFCFNKGGC 264

QY 337 PORCVNTQGGFECHCYPNYDL--VDGECVEPDPFCFRAN--CEYQCPPLNQTSLYLCVCAEGF 394  
DB 265 QHCHTNHRAQCQPGPGFHLSDRRSCVDIDECAKNGCEHCFCNVKGT--YRCKREGY 323

QY 395 APIPEPHRCQMF-----CNQATACADC--DPNTQASCECEGYIL--DGFICTDIDEC- 445  
DB 324 -QLGRDGRTEEMGLGCQVNGCGCHDQYDQPDGHHVCKRNGYILANDQKLCHDINECH 382

QY 446 ENGFGCSGVCHNLPGTFECICGPDSSALARIHGT-----DCDS-----GKVDGDSG 491  
DB 383 ENNGDCSCICVNLGAGSVECCQCKPGFRLMKDRKTCEDISECSNNGGCEQICSNQEGGYMC 442

QY 492 SGEPE 495  
DB 443 SCEP 446

RESULT 8  
O88281  
ID O88281 PRELIMINARY; PRT; 1574 AA.  
AC O88281;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE MEGF6.  
DE MEGF6.  
GN MEGF6.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;  
RX MEDLINE=98360089; PubMed=9693030;  
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;  
RT "Identification of high-molecular-weight proteins with multiple EGF-  
RT like motifs by motif-trap screening."  
RL Genomics 51:27-34(1998).  
DR EMBL; AB011532; BAA32462.1; -  
DR HSSP; P00736; IAPQ.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR Pfam; PF00008; EGF; 24.  
DR SMART; SM00179; EGF\_CA; 4.  
DR SMART; SM00001; EGF\_like; 19.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 5.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_23.  
DR PROSITE; PS01186; EGF\_2; 23.  
DR PROSITE; PS01187; EGF\_CA; 5.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
SQ SEQUENCE 1574 AA; 165445 MW; 2B48533D8F77F6E7 CRC64;

Query Match 14.2%; Score 414; DB 11; Length 1574;  
Best Local Similarity 34.1%; Pred. No. 1.4e-24;  
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 17;

QY 224 CTAPPG-AVOGHWAREAPAGWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTA 282  
DB 147 CRCPGPGYLGQD-GKTCQDVDECAHNGCGCHRCVNTPGSYLCEKPGFRLHTDGRCTL- 204

QY 283 SATQSC---NDLCEHFCVP-----NPDQPGSYSCM----- 309  
DB 205 -ATSSCTLNGGCGCHQCVLTVTHRCQCRPOYQLOEDGRRCVRSRSCAENGSGCMHLCQ 263

QY 310 -----CETGYRLAADQHRCEVDVDDCILPEPSPQRCVNTQGGFECHCYPNYDL--VD 359  
DB 264 ELRGLAHGCHGCHPGYQLAADRKTCEDVDECAALGQAQAHGCLNTQGSFKVCYCHAGELGAD 323

QY 360 G-EC---VFPVDPFCFRAN--CEYQCPPLNQTSLYLCVCAEGFAPIPHPHRCQMF---- 408  
DB 324 GRCQYRTEMEIVNSCEAGNGCGCHGCSH-TSTGPLCTCPRGY-----ELDEDQKTCIDID 377

QY 409 ---NOTACPADCPNTQA--SCECPGYILD-DGFICTDIDECNG-GFCSGVCHNLPGT 461  
DB 378 DCANSPCCQAC-ANTPGYECSCFAGYRLNTDGCCEVDDECASGHGCGEHCNLAGS 436

QY 462 FECIC 466  
DB 437 FQCF 441

RESULT 9  
 Q9V589 PRELIMINARY; PRT; 1394 AA.  
 AC Q9V589;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE CG7526 protein (fragment).  
 GN CG7526.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Bortchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT The genome sequence of *Drosophila melanogaster*.  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003558; AAF50538.1; .  
 DR HSSP; P00736; IAPQ.  
 DR FlyBase; FBgn0035798; CG7526.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF-Ca.  
 DR InterPro; IPR003410; Hyalin.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro; IPR001491; Thrbomodulin.  
 DR Pfam; PF00008; EGF; 11.  
 DR Pfam; PF00084; sushi; 2.  
 DR PRINTS; PR00907; THRBOMODULIN.  
 DR SMART; SM00032; CCP; 2.  
 DR SMART; SM00179; EGF\_CA; 9.  
 DR SMART; SM00001; EGF\_Like; 5.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 8.

DR PROSITE; PS01186; EGF\_2; 9.  
 DR PROSITE; PS01187; EGF\_CA; 10.  
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 FT NON\_TER 1394 1394  
 SQ SEQUENCE 1394 AA; 152269 MW; CD29380E3162F68A CRC64;  
 Query Match 13.2%; Score 386; DB 5; Length 1394;  
 Best Local Similarity 32.1%; Pred. No. 2.2e-22;  
 Matches 99; Conservative 33; Mismatches 110; Indels 56; Gaps 18;  
 QY 224 CTAPPGAVQG-----HWREAPGAW-----DCSVNGGCEHACNALPG 261  
 DB 563 CLCPGGYALNDHNHIVTSLNSSFTDSTSETPSAHTCLDIDECSLANGNSHFQNEPG 622  
 QY 262 APRCQCPAGALQADGRSCTASATQSC---NDLCEHFVCPNDPQSGVSCMETGYRLAA 318  
 DB 623 GFQCACPLGYALSEDMRTC--QDIDECILDSNGQCSQLCL---NORGGFACACETGTFELTP 677  
 QY 319 DQHRCEVDVDCILEPSPQPCRVNTQGGFECYCPNYDLVDGE--CVPEVDPC---FRAN 373  
 DB 678 DFGCADIDECSDYNGNSDIDICINLLGTHACACERGYELAKDKLSCLD-VDEAGLLSGG 736  
 QY 374 CEYQCQPLNQT-SYLCVCAEGFAPIPHEPHRCQFCNQTACPA--DCDPTQASC----- 425  
 DB 737 CSHEC--INKAGTECGCPLGY--ILNDGGR-----SCSPALVGCPTGQRSADGCA 785  
 QY 426 -ECPEGYLLDGFCTDIDEC--ENGPGSCGVCHNLPGTFECICGP-----DSALARHI 476  
 DB 786 IECNPGYTLGSDDKCDVIDEQKONGG--CSHRCSNTEGSKCPGPGYELDSQDKTQCDI 844  
 QY 477 GTDCDSGK 484  
 DB 845 -DECDQDK 851  
 RESULT 10  
 P87363 PRELIMINARY; PRT; 708 AA.  
 AC P87363;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Fibrillin-1 (Fragment).  
 GN FBN1.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20152896; PubMed=10691037;  
 RA Zhou G., Price C.E., Rosenquist T.H., Gadsen P.F., Godfrey M.;  
 RT Partial cloning and sequencing of chick fibrillin-1 cDNA.;  
 RL In Vitro Cell. Dev. Biol. Anim. 36:19-25(2000).  
 DR EMBL; U88872; AAB48531.1; .  
 DR HSSP; P07204; 2ADX.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF-Ca.  
 DR InterPro; IPR002212; Fibril-assoc.  
 DR InterPro; IPR001491; Thrbomodulin.  
 DR Pfam; PF00008; EGF; 13.  
 DR Pfam; PF00683; TB; 2.  
 DR PRINTS; PR00907; THRBOMODULIN.  
 DR SMART; SM00179; EGF\_CA; 14.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 13.  
 DR PROSITE; PS01186; EGF\_2; 10.  
 DR PROSITE; PS01187; EGF\_CA; 13.  
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 FT NON\_TER 708 708  
 SQ SEQUENCE 708 AA; 76164 MW; C247271C1DF73361 CRC64;

```
Query Match 12.9%; Score 376; DB 13; Length 708;
Best Local Similarity 31.8%; Pred. No. 5.6e-22;
Matches 112; Conservative 26; Mismatches 108; Indels 106; Gaps 20;

Qy 244 DCSVENGCCEHACNAIPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFVCPNP--- 300
Db 223 EGSIMNGCCNEFCCTSGSEYSCSKQGFALMPDHTCT-----DIDE--CEDPNIC 272

Qy 301 -----DQGSYSYSCMCTGYRLAADQHRCEVDVDDCILEPSPC-PQRCVNTQGGFECHCYP 353
Db 273 DGGQCTNIPGEYRCICYDGFMAEDMKTCDVNECDLHPNCLSGTCENTKGSFICHCDM 332

Qy 354 NYDLVDGE--CVEPVDPD--FRANCEYQCOPLN-QTSYLCVCAEGFA-----PIP 398
Db 333 GYSYSGKGTGCTD-INCEIGAHCNDRHVAVCTNIPGSKSCSSGWIIGNIRKTDLDECS 391

Qy 399 HEPHRCQFCNCTACPADCPNTOAS--CECPGYILDDGFICTDIDE-----CENG- 448
Db 392 NGTHKCSPH-----ADC-KNTMGSYRLCKEGY-TGDGFTCTDLDCESENINLCENQ 442

Qy 449 -----GF-----CS-----GVCHNLPCTFEICGPDPSALA 473
Db 443 CLNAPGYRCECDMGFLPSLDKACEDIDCSLPNICVYGTCHNLPGLFRCEVEGYELD 502

Qy 474 RHIG-----TDCSGK-VDGGDSGSGEPSPPTPGSTLTPPAVGLV 513
Db 503 RSGNGCTDNECADPTTCISGTCVNTAGSYTCECPDP-----PELNPTRVGCV 550

RESULT 11
Q9WUH9
ID Q9WUH9 PRELIMINARY; PRT; 2906 AA.
AC Q9WUH9
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Fibrillin-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99350231; PubMed=10419698;
RA Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashiwara N., Wallner E.,
RA Kanwar Y.S.;
RT "Cloning of rat fibrillin-2 cDNA and its role in branching
RT morphogenesis of embryonic lung.";
RL Dev. Biol. 212:229-242(1999).
DR EMBL; AF135060; AAD34439.1; -
DR HSSP; P35555; IEMN.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002212; Fibril-assoc.
DR Pfam; PF00008; EGF; 46.
DR PRINTS; PR00010; EGFHLOOD.
DR SMART; SM00179; EGF_CA; 42.
DR SMART; SM00001; EGF_like; 4.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 43.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2906 AA; 313371 MW; 9EE64E727044EF58 CRC64;

Query Match 12.9%; Score 375; DB 11; Length 2906;
Best Local Similarity 32.9%; Pred. No. 3.8e-21;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;
```

```
Qy 244 DCSVENGCCEHACNAIPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFVCPNP-- 301
Db 1237 ECMIMNGGCDTQCTNSEGSEYSCSGYALMPGRSCA-----DIDE--CENNPDI 1286

Qy 302 -----QPGSYSCMCTGYRLAADQHRCEVDVDDCILEPSPC-PQRCVNTQGGFECHCYP 353
Db 1287 DGGQCTNIPGEYRCICYDGFMAEDMKTCDVNECDLNPNCMFEGECENTKGSFICHQ 1346

Qy 354 NYDLVDGE--CVEPVDPD--FRANCEYQCOPLN-QTSYLCVCAEGFA-----PIP 398
Db 1347 GYSVKKGATGCTD-VDECEIGAHCNCDMHASCLNVPGSKSCREGWVGNGIKIDLDECA 1405

Qy 399 HEPHRCQFCNCTACPADCPNTOAS--CECPGYILDDGFICTDIDE-----CENG- 448
Db 1406 NGTHQCSI-----NAQC-VNTPGYSYRCACSEGF-TGDGFTCTSDVDECAENINLCENQ 1456

Qy 449 -----GF-----CS-----GVCHNLPCTFEICGPDPSALA 473
Db 1457 CLNAPGYRCECEMGFTPASDSRSCQDIDCSFQNICVFTGTCNNLPGMFHCICDDGYGLD 1516

Qy 474 RHIG--TDGD 481
Db 1517 RTGGHCTDID 1526

RESULT 12
O75095
ID O75095 PRELIMINARY; PRT; 1246 AA.
AC O75095
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MEGF6 (Fragment).
GN MEGF6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98360089; PubMed=9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RT like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
DR EMBL; AB011539; BAA32467.2; -
FT NON_TER
SQ SEQUENCE 1246 AA; 130304 MW; BD8E70A0F6A2CB29 CRC64;

Query Match 12.7%; Score 369; DB 4; Length 1246;
Best Local Similarity 32.6%; Pred. No. 4.5e-21;
Matches 95; Conservative 27; Mismatches 97; Indels 72; Gaps 15;

Qy 238 EAPGAWDCSVENGCCEHACNAIPGAPRCOCAGALQADGRSCTASATQSC---NDLCEH 294
Db 70 QRPDVDECTHNGGCGHRCVNTPGSYLCEKPGFRLHDSRTCL--AINSCALNGGCGH 127

Qy 295 FCV-----PNDQPGS-----YSCMCTGYR 315
Db 128 HCVQLTITRHCQCRPGFQLQEDGRHCVRRSPCANRNGSCMHRQCVVRGLARCECHVGQ 187

Qy 316 LAADQHRCEVDVDDCILEPSPCQRCVNTQGGFECHCYPNIDL-VDG-EC-----VEPVDP 369
Db 188 LAADGKACEDVDECAAGLAQCAHGLCNTQGSFKVCVCHAGYELGADGRQCYRIEMEIVNSC 247

Qy 370 FRAN--CEYQCOPLNQTSY--LCVCAEGFAPIPHEPHRCQFC-----NOTACPADCD 418
Db 248 EANNNGCASHGC---SHTSAGPLCTCPRGY-----ELDTQRTCIDVDCCADSPCCQOVCT 299

Qy 419 PNTQA-SCCEPBGYILD-DGFICTDIDECENG-GFCSGVCHNLPCTFEIC 466
```



Db 300 NNPGGYECGYAGYRLSADGGCGEDYDECASSRGGEHHCNTLAGSFQCS 350

RESULT 13  
O88840

```
ID O88840 PRELIMINARY; PRT; 3857 AA.
AC O88840;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mutant fibrillin-1.
GN FBN1
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP "Structure of the mutant fibrillin-1 gene in the tight skin (TSK)
RC mouse."
RL DNA Res. 4:267-271(1997).
DR EMBL; AF007248; AAC62317.1; -.
DR HSP; P35555; IAPU.
DR MGD; MGI:95489; Fbn1.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR002557; Chitin_bind_perA.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002212; Fibril-assoc.
DR InterPro; IPR000822; znf_C2H2.
DR Pfam; PF00008; EGF; 64.
DR Pfam; PF00683; TB; 12.
DR SMART; SM00494; ChtBD2; 2.
DR SMART; SM00179; EGF_Ca; 60.
DR SMART; SM00001; EGF_like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 61.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 50.
DR PROSITE; PS01187; EGF_Ca; 61.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 3857 AA; 418301 MW; 58C0618BC527E04C CRC64;
```

```
Query Match 12.6%; Score 367; DB 11; Length 3857;
Best Local Similarity 30.3%; Pred. No. 2.2e-20;
Matches 118; Conservative 28; Mismatches 132; Indels 112; Gaps 21;

QY 212 SAAVAPLGL-----QLMCTAPPAGVQGHWAREAPGAWDCSVENGGEHACNAIPGAPRC 265
Db 2148 SANLCPHRCVNLGKQACNFGYHPTHDLFCVDIDECISNMGCGEFTCTNSDGSYEC 2207

QY 266 QCPAGALQADGRSCTASATQSCNDLCEHFCVNP-----DQPGYSVCMTGYRL 316
Db 2208 SCQGFALPDRSCT-----DIDE--CEDPNICDGGQCTNIPGEYRCICYDGFMA 2257

QY 317 AADQHRCEYDDCIIEPSPC-PQRCVNTGGFECYCPNYDLVDGE--CVEPYDPC--FR 371
Db 2258 SEDMKTCDVNECDLNPICLSGTCTKNTGSKFICHDMGYSKKGTGTCTD-INECEIGA 2316

QY 372 ANCYOCQPLNOT-SYLCVCAEGFA-----PIPEPHRCQMFNCQTACPADCPN 420
Db 2317 HNCDRHAVCTNTAGSFKSCSPGWIGDKIKCTDLDECSSGTHMCSQH-----ADC-KN 2368

QY 421 TQAS--CECEPEGLDDGFTCTDIDEC-EN-----GGF----- 450
Db 2369 TMGSYRCLCKDGY-TGDGFTCTDLDECSENLCNGQCILNAPGGYRCBCDMGFVPSADG 2427

QY 451 -----CS-----GVCHNLPGTFECICGPDPSALARIHG-----TDCDSGK 484
```

Db 2428 KACEDIDECSLPNICVFCGTCHNLPGFPRCECEIGVELDRSGGNCCTDYNECLDPTTCISGN 2487

QY 485 -VDGDSGSGEPFPPSPTPGSTLTTPAVGLV 513

Db 2488 CVNTPGSGYTCDCPPD---FELNPTRVGCV 2513

## RESULT 14

```
Q96JB8
ID Q96JB8 PRELIMINARY; PRT; 2809 AA.
AC Q96JB8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Fibrillin3.
GN KIAA1776.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 8:85-95(2001).
DR EMBL; AB053450; BAB47408.1; -.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002212; Fibril-assoc.
DR Pfam; PF00008; EGF; 45.
DR Pfam; PF00683; TB; 9.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_41.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_36.
DR PROSITE; PS01187; EGF_Ca; UNKNOWN_40.
DR PROSITE; PS01187; EGF_Ca; UNKNOWN_40.
SQ SEQUENCE 2809 AA; 300323 MW; 20C04CC006C0161F CRC64;
```

```
Query Match 12.5%; Score 365; DB 4; Length 2809;
Best Local Similarity 30.0%; Pred. No. 2.3e-20;
Matches 107; Conservative 27; Mismatches 107; Indels 116; Gaps 20;

QY 244 DCSVENGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP--- 300
Db 1158 ECRVQNGGCDVHCINTGYSYRCSCGGYSLMPDGRACA-----DYDE--CEENPRVC 1207

QY 301 -----DQPGYSVCMTGYRLAADQHRCEYDDCIIEPSPCPO-RCVNTGGGFFCHCYP 353
Db 1208 DQGHCTNMPGGHRCICYDGFMATPMRTCDVDDECDLNPICLHGDGENTKGSFVCHQCL 1267

QY 354 NYDLVDGE--CVEPYDPC-----FRANC----- 374
Db 1268 GTMWRKATGCSGSD-VDECEVGHNCDSHASCINIPGSPSCRLPGWVGDFECHDLDCI 1326

QY 375 --EYQCP-----LN-QTSYLCVCAEGFAPIPHEPHRCQMF--NQTACPADCD--PNTQA- 423
Db 1327 SQEHRCSRPGDCLNVPYSYRCTGROFA-----GDGFFCEDRDECAENVLDLDCNGOCL 1379

QY 424 -----SECEPGY-ILDDGFTCTDIDECENGFFGS-GVCHNLPGTFECICGPDPSALARIH 475
Db 1380 NAGPYRCBCEMGFDPTEDHRACQDWDCAQNLCAFGSCENLPGMFRICNG----- 1432

QY 476 ICTDCDSKGVDSGSGSEPPSP-----TPGS-----TLPVAVGLV 513
Db 1433 -GYELDRG---GGNCTDINECADPVNCGINGVINTPGSYLCSGPDGFELNPGVGCV 1485

RESULT 15
Q9WUH8
```



Search completed: December 16, 2002, 17:26:17  
Job time : 40.5 secs

ID Q9WUH8 PRELIMINARY; PRT; 2872 AA.  
AC Q9WUH8;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Fibrillin-1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID-10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-95032689; PubMed-9815129;  
RA Kanwar Y.S., Ota K., Yang Q., Kumar A., Wada J., Kashiwara N.,  
RA Peterson D.R.;  
RT \*Isolation of rat fibrillin-1 cDNA and its relevance in metanephric  
RT development.\*;  
RL Am. J. Physiol. 275:F710-F723(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kanwar Y.S.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF135059; AAD34438.1; -.  
DR HSP; P35555; 1APJ.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF Ca.  
DR InterPro; IPR002212; Fibril-assoc.  
DR InterPro; IPR000822; Znf\_C2H2.  
DR Pfam; PF00008; EGF; 45.  
DR Pfam; PF00683; TB; 9.  
DR SMART; SM00179; EGF\_CA; 41.  
DR SMART; SM00001; EGF\_like; 5.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 42.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_2.  
DR PROSITE; PS01186; EGF\_2; 38.  
DR PROSITE; PS01187; EGF\_CA; 41.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
SQ SEQUENCE 2872 AA; 312069 MW; 0C4F7F3B87A80280 CRC64;  
  
Query Match 12.4%; Score 361; DB 11; Length 2872;  
Best Local Similarity 30.6%; Pred. NO. 4.9e-20;  
Matches 114; Conservative 25; Mismatches 127; Indels 106; Caps 20;  
  
QY 224 CTAPPGAVQGHWAEPAGWDCSVENGCGEHCACNAIPGAPRCOCPAGAAQAQAGRSCTAS 283  
DB 1181 CACNPGYHPTHDLFCVDIDECSIMNGCETFTCTNSDGYECSCQPGFALMPQORSET-- 1238  
QY 284 ATQSCNDLCHEFCVPNP-----DQPGSYSCMCETGYRLAADOHRCEVDVDDCILEPS 334  
DB 1239 -----DIDE--CEDNPNICDGGQCTNIPGEYRCLCYDGFWASEDMKTCVDVNECDLNP 1290  
QY 335 PC-PQRCVNTQGGFECHCPNYDLVDGE--CPEVPDPC--FRANCEYQCQPLNOT-SYLC 388  
DB 1291 ICLSGTCENTKGFICHCDMGYSCKKCKTGCTD-INECEIGHNCGRHAVCTNTAGSFKC 1349  
QY 389 VCAEGFA-----PIPEPHRCOMFNQACPADCDPNTQAS--CECPGYILDDG 436  
DB 1350 SCSPGWIGDGIKCTDLDGCSNGTHMCSOH-----ADC-KNTMGSYRCLCKDGY-TGDG 1400  
QY 437 FICTDIDECE-----NG-----GF-----CS-----G 453  
DB 1401 FTCTDLDECSENLSNGQCLNAPAGYRCEDMGFVPSADGKACEDINECSLPNICVFG 1460  
QY 454 VCHNLPGTFECICGPDSSALARHIG-----TDCDSCK-VDGGDSGSGEPPSPPTP 501  
DB 1461 TCHNLPLGLFRCIEIGYELDRSGGNCVDVNECLDPTTTCISGNCVNTPGSYTCDPCPD--- 1517  
QY 502 GSTLTTPPAVGLV 513  
DB 1518 -FELNPTRVGCV 1528



Result No.	Query %			ID	Description
	Score	Match	Length		
1	2916	100.0	575	1	US-08-312-870-1
2	2916	100.0	575	6	5466668-6
3	2912	99.9	575	1	US-08-261-206A-59
4	2908	99.7	575	1	US-08-170-290A-54
5	2846.5	97.6	572	6	5256770-7
6	2826	96.9	498	2	US-08-733-564-2
7	2824	96.8	497	1	US-08-312-870-3
8	2766	94.9	494	1	US-08-014-723-14
9	2766	94.9	494	1	US-08-110-011A-14
10	2764	94.8	494	1	US-08-014-723-16
11	2764	94.8	494	1	US-08-110-011A-16
12	2690	92.2	475	1	US-08-307-444A-2
13	2690	92.2	475	1	US-08-587-389-2
14	2886	92.1	475	1	US-08-307-444A-1
15	2886	92.1	475	1	US-08-587-389-1
16	2680	91.9	476	1	US-08-014-723-1
17	2680	91.9	476	1	US-08-110-011A-1
18	2678	91.8	476	1	US-08-014-723-2
19	2678	91.8	476	1	US-08-014-723-18
20	2678	91.8	476	1	US-08-110-011A-2
21	2678	91.8	476	1	US-08-110-011A-18
22	2596	89.0	456	1	US-08-307-444A-4
23	2596	89.0	456	1	US-08-587-389-4
24	2592	88.9	456	1	US-08-307-444A-3
25	2592	88.9	456	1	US-08-587-389-3
26	2543	87.2	446	1	US-08-307-444A-5
27	2543	87.2	446	1	US-08-587-389-5

QY 1 MLGVLVILGALALAGLGFPAPEPQPGSQCVHDCFCALYPGPATFLNASQICDGLRGLM 60  
Db 1 MLGVLVILGALALAGLGFPAPEPQPGSQCVHDCFCALYPGPATFLNASQICDGLRGLM 60  
QY 61 TVRSSVAADVLSLLNGDGVGRRRLWIGLQPPGCGDKRGLGPGFQWVTDGNNNTSYS 120  
Db 61 TVRSSVAADVLSLLNGDGVGRRRLWIGLQPPGCGDKRGLGPGFQWVTDGNNNTSYS 120  
QY 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFCFHFHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFCFHFHFPATCRPLAV 180  
QY 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
Db 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCEYRLLAADQHRCEYVDCILEPSPQRCVNTQGGFECCHYFNVDLVDG 360  
Db 301 DQPGSYSCMCEYRLLAADQHRCEYVDCILEPSPQRCVNTQGGFECCHYFNVDLVDG 360  
QY 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQACPADCDPN 420  
Db 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQACPADCDPN 420  
QY 421 TOASCECEGYILDDGFCITDIDECENGFCGVCVCHNLPGTFECICGPDSSALARIHIGTDC 480  
Db 421 TOASCECEGYILDDGFCITDIDECENGFCGVCVCHNLPGTFECICGPDSSALARIHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
RESULT 2  
546668-6  
; Patent No. 546668  
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,  
; DAVID R.  
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR  
; PHARMACEUTICAL USE  
; NUMBER OF SEQUENCES: 57  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/155,346  
; FILING DATE: 22-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 568,456  
; FILING DATE: 15-AUG-1990  
; APPLICATION NUMBER: 506,325  
; FILING DATE: 09-APR-1990  
; APPLICATION NUMBER: 406,941  
; FILING DATE: 13-SEP-1989  
; APPLICATION NUMBER: 345,374  
; FILING DATE: 28-APR-1989  
; SEQ ID NO: 6  
; LENGTH: 575  
546668-6

Query Match 100.0%; Score 2916; DB 6; Length 575;  
Best Local Similarity 100.0%; Pred. No. 8.3e-199; Indels 0; Gaps 0;  
Matches 516; Conservative 0; Mismatches 0;

QY 1 MLGVLVILGALALAGLGFPAPEPQPGSQCVHDCFCALYPGPATFLNASQICDGLRGLM 60  
Db 1 MLGVLVILGALALAGLGFPAPEPQPGSQCVHDCFCALYPGPATFLNASQICDGLRGLM 60  
QY 61 TVRSSVAADVLSLLNGDGVGRRRLWIGLQPPGCGDKRGLGPGFQWVTDGNNNTSYS 120  
Db 61 TVRSSVAADVLSLLNGDGVGRRRLWIGLQPPGCGDKRGLGPGFQWVTDGNNNTSYS 120

QY 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFCFHFHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFCFHFHFPATCRPLAV 180  
QY 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
Db 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCEYRLLAADQHRCEYVDCILEPSPQRCVNTQGGFECCHYFNVDLVDG 360  
Db 301 DQPGSYSCMCEYRLLAADQHRCEYVDCILEPSPQRCVNTQGGFECCHYFNVDLVDG 360  
QY 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQACPADCDPN 420  
Db 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQACPADCDPN 420  
QY 421 TOASCECEGYILDDGFCITDIDECENGFCGVCVCHNLPGTFECICGPDSSALARIHIGTDC 480  
Db 421 TOASCECEGYILDDGFCITDIDECENGFCGVCVCHNLPGTFECICGPDSSALARIHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
RESULT 3  
US-08-261-206A-59  
; Sequence 59, Application US/08261206A  
; Patent No. 5574007  
; GENERAL INFORMATION:  
; APPLICANT: Gomi, Mitichitaka  
; APPLICANT: Zushi, Komakazu  
; APPLICANT: Yamamoto, Shuji  
; APPLICANT: Suzuki, Koji  
; APPLICANT: Matsuda, Akio  
; TITLE OF INVENTION: A Polypeptide Capable of Interacting  
; WITH THROMBIN  
; NUMBER OF SEQUENCES: 80  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: 301 N. Washington St.  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22046-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/261,206A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/740,492  
; FILING DATE: 03-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svensson, Leonard R.  
; REGISTRATION NUMBER: 30330  
; REFERENCE/DOCKET NUMBER: 216-275P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-241-1300  
; TELEFAX: 703-241-2846  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 575 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..575  
OTHER INFORMATION: /label= protein  
OTHER INFORMATION: /note= "human thrombomodulin"  
US-08-261-206A-59

Query Match 99.9%; Score 2912; DB 1; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.6e-198;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGVLVLGALALAGLGFPAEAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MGVLVLGALALAGLGFPAEAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQPPCGDPPKRLGPGFOWVTGDNNTSYS 120  
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQPPCGDPPKRLGPGFOWVTGDNNTSYS 120

Qy 121 RWARDLNGAPLCGPLCVAVSAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180  
Db 121 RWARDLNGAPLCGPLCVAVSAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180

Qy 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVGHHAREAP 240  
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVGHHAREAP 240

Qy 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVPNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVPNP 300

Qy 301 DQGSYSVCMECYGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
Db 301 DQGSYSVCMECYGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360

Qy 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
Db 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

Qy 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
Db 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480

Qy 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHSG 516

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/170,290A  
FILING DATE: 28-DEC-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/05573  
FILING DATE: 01-JUL-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/724,237  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Heslin, James M.  
REGISTRATION NUMBER: 29,541  
REFERENCE/DOCKET NUMBER: 11972-58-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 575 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-170-290A-54

Query Match 99.7%; Score 2908; DB 1; Length 575;  
Best Local Similarity 99.8%; Pred. No. 3.1e-198;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGVLVLGALALAGLGFPAEAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MGVLVLGALALAGLGFPAEAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQPPCGDPPKRLGPGFOWVTGDNNTSYS 120  
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQPPCGDPPKRLGPGFOWVTGDNNTSYS 120

Qy 121 RWARDLNGAPLCGPLCVAVSAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180  
Db 121 RWARDLNGAPLCGPLCVAVSAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180

Qy 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVGHHAREAP 240  
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVGHHAREAP 240

Qy 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVPNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVPNP 300

Qy 301 DQGSYSVCMECYGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
Db 301 DQGSYSVCMECYGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360

Qy 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
Db 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

Qy 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
Db 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480

Qy 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHSG 516

RESULT 5

```
5256770-7
; Patent No. 5256770
; APPLICANT: GLASER, CHARLES B.; MORSER, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/506,325
; FILING DATE: 09-APR-1990
; SEQ ID NO: 7
; LENGTH: 572
5256770-7

Query Match          97.6%; Score 2846.5; DB 6; Length 572;
Best Local Similarity 98.6%; Pred. No. 6.8e-194;
Matches 509; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 1 MLGVVLGALALAGLGFPAEPPQPGSQCVHDCFALYPGATFLNASQICDGLRHLM 60
DB 1 MLGVVLGALALAGLGFPAEPPQPGSQCVHDCFALYPGATFLNASQICDGLRHLM 60
QY 61 TVRSSVAADVLSLLNGDGVGRRRLWIGLQPLPPGCGDPKRLGFLRGFWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGVG--RRLWIGLQPLPPGCGDPKRLGFLRGFWVTGDNNTSYS 119
QY 121 RWARLDUNGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
DB 121 RWARLDUNGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 179
QY 181 EPGAAAVSITVTGTPFAARGDFOALPGVSSAAVAPLGLQMLCTAPPAGVGHWAREAP 240
DB 180 EPGAAAVSITVTGTPFAARGDFOALPGVSSAAVAPLGLQMLCTA--GNVOGHWAREAP 237
QY 241 GAWDCSVENGCGCEHACNAPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 238 GAWDCSVENGCGCEHACNAPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 297
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNVDLVG 360
DB 298 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNVDLVG 357
QY 361 ECVEPVPDPCFRANCEYQCPNLTGYLCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420
DB 358 ECVEPVPDPCFRANCEYQCPNLTGYLCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 417
QY 421 TQASCEPEGYILDDGFICTDIDECENGCGFCGVCHNLPFTFECIGPDSALARHIGTDC 480
DB 418 TQASCEPEGYILDDGFICTDIDECENGCGFCGVCHNLPFTFECIGPDSALARHIGTDC 477
QY 481 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 516
DB 478 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 513

RESULT 6
US-08-733-564-2
; Sequence 2, Application US/08733564
; Patent No. 5916874
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, Kenji
; APPLICANT: MOCHIDA, Satoshi
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 2
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/733,564
FILING DATE: 18 OCTOBER 1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, Leonard R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 0216-0362P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-733-564-2

Query Match          96.9%; Score 2826; DB 2; Length 498;
Best Local Similarity 99.8%; Pred. No. 1.7e-192;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 APAEPQPGSQCVHDCFALYPGATFLNASQICDGLRHLMTVRSSVAADVLSLLNGD 78
DB 1 APAEPQPGSQCVHDCFALYPGATFLNASQICDGLRHLMTVRSSVAADVLSLLNGD 60
QY 79 GGVRRRLWIGLQPLPPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDUNGAPLCGPLCV 138
DB 61 GGVRRRLWIGLQPLPPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDUNGAPLCGPLCV 120
QY 139 AVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIITGTPFA 198
DB 121 AVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIITGTPFA 180
QY 199 ARGADFOALPGVSSAAVAPLGLQMLCTAPPAGVGHWAREAPGAWDCSVENGCGCEHACNA 258
DB 181 ARGADFOALPGVSSAAVAPLGLQMLCTAPPAGVGHWAREAPGAWDCSVENGCGCEHACNA 240
QY 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYSCMETGYRLAA 318
DB 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYSCMETGYRLAA 300
QY 319 DQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNVDLVGECVPEVPDPCFRANCEYQC 378
DB 301 DQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNVDLVGECVPEVPDPCFRANCEYQC 360
QY 379 QPLNQTGYLCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCEPEGYILDDGFI 438
DB 361 QPLNQTGYLCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCEPEGYILDDGFI 420
QY 439 CTDIDECENGCGFCGVCHNLPFTFECIGPDSALARHIGTDCDQSGKVDGSDSGSEPPSP 498
DB 421 CTDIDECENGCGFCGVCHNLPFTFECIGPDSALVRHIGTDCDQSGKVDGSDSGSEPPSP 480
QY 499 PTPGSTLTPPAVGLVHSG 516
DB 481 PTPGSTLTPPAVGLVHSG 498

RESULT 7
US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-312-870-3

Query Match          96.8%; Score 2824; DB 1; Length 497;
Best Local Similarity 100.0%; Pred. No. 2.3e-192;
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAEPQGGSCVDEHDFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 78
Db 1 APAEPQGGSCVDEHDFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 60
QY 79 GVGRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 138
Db 61 GVGRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
QY 139 AVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAAASITYGTTPFA 198
Db 121 AVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAAASITYGTTPFA 180
QY 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGCEHACNA 258
Db 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGCEHACNA 240
QY 259 IPGAPRCQCPAGALQADGRSCTASATOSCNLDLCEHFCVPPNPDPQPGSYSCMCETGYRLAA 318
Db 241 IPGAPRCQCPAGALQADGRSCTASATOSCNLDLCEHFCVPPNPDPQPGSYSCMCETGYRLAA 300
QY 319 DOHRCEDVDDCILEPSPQRCVNTQGGFECYCHYNYDLVDGECVPEVDPDPCFRANCEYOC 378
Db 301 DOHRCEDVDDCILEPSPQRCVNTQGGFECYCHYNYDLVDGECVPEVDPDPCFRANCEYOC 360
QY 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQNTACPADCPNTQASCECEGYILLDDGFI 438
Db 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQNTACPADCPNTQASCECEGYILLDDGFI 420
QY 439 CTDIDECENGFGCSGVCHNLPGTFECICGPDSSALARIHGTDCDSKVGDDSGSGSEPPPS 498
Db 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDSSALARIHGTDCDSKVGDDSGSGSEPPPS 480
QY 499 PTPGSTLTTPPAGVLVHS 515
Db 481 PTPGSTLTTPPAGVLVHS 497

RESULT 8
```

```
US-08-014-723-14
; Sequence 14, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-014-723-14

Query Match          94.9%; Score 2766; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.9e-188;
Matches 490; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFAPAEPQPGSQCVDEHDFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MLGVLVLGALALAGLGFAPAEPQPGSQCVDEHDFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVGRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSY 120
Db 61 TVRSSVAADVISLLNGDGGVGRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSY 120
QY 121 RWARLDLNGAPLCGLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
Db 121 RWARLDLNGAPLCGLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
QY 181 EPGAAAAAASITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAAAASITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGALQADGRSCTASATOSCNLDLCEHFCVPPNP 300
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGALQADGRSCTASATOSCNLDLCEHFCVPPNP 300
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
QY 361 ECVEPVDPDPCFRANCEYOCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQNTACPADCPN 420
```

|||||  
Db 361 ECVEPVDPCFRANCEYQCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420  
Qy 421 TQASCEPEGILDGFGICTDIDECENGFGCGVCHNLPGTFECICGPDSSALARHIGTDC 480  
Db 421 TQASCEPEGILDGFGICTDIDECENGFGCGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Qy 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDDEDYSGSGE 494  
  
RESULT 9  
US-08-110-011A-14  
; Sequence 14, Application US/08110011A  
; Patent No. 5354664  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,011A  
; FILING DATE: 23-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5354664man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-110-011A-14  
  
Query Match 94.9%; Score 2766; DB 1; Length 494;  
Best Local Similarity 99.2%; Pred. No. 2.9e-188;  
Matches 490; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 1 MLGVLVLGALALAGLGFAPAEPPQGGSCVCHDCEFCALYPGPATFLNASQICDGLRGLHLM 60  
Db 1 MLGVLVLGALALAGLGFAPAEPPQGGSCVCHDCEFCALYPGPATFLNASQICDGLRGLHLM 60  
Qy 61 TVRSSVAADVVISLLNGDCGVRRLRWIGLQLPCCGDKPKRLGPLRGFWGTGDNNTSYS 120  
Db 61 TVRSSVAADVVISLLNGDCGVRRLRWIGLQLPCCGDKPKRLGPLRGFWGTGDNNTSYS 120  
Qy 121 RWARLDNGLAPLCGLPLCAVSAAEATVPSEPTWEEQQCEVKADGFTLCEHFFPATCPLAV 180  
Db 121 RWARLDNGLAPLCGLPLCAVSAAEATVPSEPTWEEQQCEVKADGFTLCEHFFPATCPLAV 180

Qy 181 EPGAAAAAVSITYCTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
Db 181 EPGAAAAAVSITYCTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
Qy 241 GAWDCSVENGCGEACRACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGEACRACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
Qy 301 DPGSYSCMCETGYRLAADOHRCEDVDCCILEPSPCPORCVNTQGGFCHCYPNYDLVDG 360  
Db 301 DPGSYSCMCETGYRLAADOHRCEDVDCCILEPSPCPORCVNTQGGFCHCYPNYDLVDG 360  
Qy 361 ECVEPVDPCFRANCEYQCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420  
Db 361 ECVEPVDPCFRANCEYQCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420  
Qy 421 TQASCEPEGILDGFGICTDIDECENGFGCGVCHNLPGTFECICGPDSSALARHIGTDC 480  
Db 421 TQASCEPEGILDGFGICTDIDECENGFGCGVCHNLPGTFECICGPDSSALVRHIGTDC 480  
Qy 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDDEDYSGSGE 494  
  
RESULT 10  
US-08-014-723-16  
; Sequence 16, Application US/08014723  
; Patent No. 5273962  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,723  
; FILING DATE: 19930208  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5273962man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-014-723-16

Query Match 94.8%; Score 2764; DB 1; Length 494;  
Best Local Similarity 99.0%; Pred. No. 4e-188;  
Matches 489; Conservative 1; Mismatches 4; Indels 0; Gaps 0;



```
QY 1 MLGVLVIGALALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MLGVLVIGALALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWRDLNLGAPLCGLPLCAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFFHPCATCRPLAV 180
Db 121 RWRDLNLGAPLCGLPLCAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFFHPCATCRPLAV 180
QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHACNATPGAPRCOCAGAAQADGRSCTASATOSCNLCEHFCVNP 300
Db 241 GAWDCSVENGCGEHACNATPGAPRCOCAGAAQADGRSCTASATOSCNLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYCNIDLV 360
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYCNIDLV 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLYCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSLYCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALRHIGTDC 480
Db 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVDDEASGSGD 494

RESULT 11
US-08-110-011A-16
; Sequence 16, Application US/08110011A
; Patent No. 5354664
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
```

```
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-16

Query Match 94.8%; Score 2764; DB 1; Length 494;
Best Local Similarity 99.0%; Pred. No. 4e-188;
Matches 489; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVIGALALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MLGVLVIGALALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWRDLNLGAPLCGLPLCAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFFHPCATCRPLAV 180
Db 121 RWRDLNLGAPLCGLPLCAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFFHPCATCRPLAV 180
QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHACNATPGAPRCOCAGAAQADGRSCTASATOSCNLCEHFCVNP 300
Db 241 GAWDCSVENGCGEHACNATPGAPRCOCAGAAQADGRSCTASATOSCNLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYCNIDLV 360
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYCNIDLV 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLYCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSLYCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALRHIGTDC 480
Db 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVDDEASGSGD 494

RESULT 12
US-08-307-444A-2
; Sequence 2, Application US/08307444A
; Patent No. 5516659
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/307,444A  
;; FILING DATE: 19-SEP-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/835,436  
;; FILING DATE: 26-FEB-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OLIFF, JAMES A.  
;; REFERENCE/DOCKET NUMBER: JAO 27706  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 836-6400  
;; TELEFAX: (703) 836-2787  
;; TELEFAX: 90-1799 PTO ALEX  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 475 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; US-08-307-444A-2

Query Match 92.2%; Score 2690; DB 1; Length 475;  
Best Local Similarity 100.0%; Pred. No. 6.5e-183;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLM 60  
DB 1 MLGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPLPGCGDKRGLPLRGFQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPLPGCGDKRGLPLRGFQWVTGDNNTSYS 120  
QY 121 RWARLDNLGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFFATCRPLAV 180  
DB 121 RWARLDNLGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFFATCRPLAV 180  
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEHFCVNP 300  
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEHFCVNP 300  
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360  
DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360  
QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
QY 421 TQASCEPEGYILDDGFTCTDIDBENGFGFCGVCNHLNPTFTFECICGPDSSALARRH 475  
DB 421 TQASCEPEGYILDDGFTCTDIDBENGFGFCGVCNHLNPTFTFECICGPDSSALARRH 475

## RESULT 13

US-08-587-389-2  
; Sequence 2, Application US/08587389  
; Patent No. 5695964  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI

;; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT  
;; TITLE OF INVENTION: PRODUCTION THEREOF, AND THERAPEUTIC AGENT (AS AMENDED)  
;; NUMBER OF SEQUENCES: 27  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OLIFF & BERRIDGE  
;; STREET: P.O. BOX 19928  
;; CITY: ALEXANDRIA  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22320  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/587,389  
;; FILING DATE: 17-JAN-1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/307,444  
;; FILING DATE: 19-SEP-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OLIFF, JAMES A.  
;; REFERENCE/DOCKET NUMBER: JAO 27706  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 836-6400  
;; TELEFAX: (703) 836-2787  
;; TELEFAX: 90-1799 PTO ALEX  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 475 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; US-08-587-389-2

Query Match 92.2%; Score 2690; DB 1; Length 475;  
Best Local Similarity 100.0%; Pred. No. 6.5e-183;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLM 60  
DB 1 MLGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPLPGCGDKRGLPLRGFQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVLSLLNGDGGVGRRLRWIGLQPLPGCGDKRGLPLRGFQWVTGDNNTSYS 120  
QY 121 RWARLDNLGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFFATCRPLAV 180  
DB 121 RWARLDNLGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFFATCRPLAV 180  
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEHFCVNP 300  
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEHFCVNP 300  
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360  
DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360  
QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TOASCCEPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFFECICGPDSSALARRH 475  
 Db 421 TOASCCEPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFFECICGPDSSALARRH 475

RESULT 14  
 US-08-307-444A-1  
 ; Sequence 1, Application US/08307444A  
 ; Patent No. 5518659  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NII, ATSUSHI  
 ; APPLICANT: MORISHITA, HIDEAKI  
 ; APPLICANT: UEMURA, AKIO  
 ; APPLICANT: MOCHIDA, EI  
 ; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OLIFF & BERRIDGE  
 ; STREET: P.O. BOX 19928  
 ; CITY: ALEXANDRIA  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22320  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/307,444A  
 ; FILING DATE: 19-SEP-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 26-FEB-1992  
 ; FILING DATE: 26-FEB-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OLIFF, JAMES A.  
 ; REGISTRATION NUMBER: 27,075  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-6400  
 ; TELEFAX: (703) 836-2787  
 ; TELEX: 90-1799 PTO ALEX  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 475 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; US-08-307-444A-1

Query Match 92.1%; Score 2686; DB 1; Length 475;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-182;  
 Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLGVLVLGALAGLGPAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRHLH 60  
 Db 1 MLGVLVLGALAGLGPAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRHLH 60  
 QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPPCGDPRKRLGFGQWVTGDNNTSYS 120  
 Db 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPPCGDPRKRLGFGQWVTGDNNTSYS 120  
 QY 121 RWARLDNGAPLCPLCVAVSAAATVPSEPIWEEQCEVKADGFLCFEFHPATCRPLAV 180  
 Db 121 RWARLDNGAPLCPLCVAVSAAATVPSEPIWEEQCEVKADGFLCFEFHPATCRPLAV 180  
 QY 181 EPGAAAAVSTYGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPVGVQGHWAREAP 240  
 Db 181 EPGAAAAVSTYGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPVGVQGHWAREAP 240

QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCEGYRLAADOHRCEVDVDCILPEPSPQRCVANTOGGFECHCYENYDLVDG 360  
 Db 301 DQPGSYSCMCEGYRLAADOHRCEVDVDCILPEPSPQRCVANTOGGFECHCYENYDLVDG 360  
 QY 361 ECVPEVPDPCFRANCEYQCOPLNQTSYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
 Db 361 ECVPEVPDPCFRANCEYQCOPLNQTSYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
 QY 421 TOASCCEPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFFECICGPDSSALARRH 475  
 Db 421 TOASCCEPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFFECICGPDSSALARRH 475

RESULT 15  
 US-08-587-389-1  
 ; Sequence 1, Application US/08587389  
 ; Patent No. 565964  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NII, ATSUSHI  
 ; APPLICANT: MORISHITA, HIDEAKI  
 ; APPLICANT: UEMURA, AKIO  
 ; APPLICANT: MOCHIDA, EI  
 ; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OLIFF & BERRIDGE  
 ; STREET: P.O. BOX 19928  
 ; CITY: ALEXANDRIA  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22320  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/587,389  
 ; FILING DATE: 17-JAN-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/307,444  
 ; FILING DATE: 19-SEP-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OLIFF, JAMES A.  
 ; REGISTRATION NUMBER: 27,075  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-6400  
 ; TELEFAX: (703) 836-2787  
 ; TELEX: 90-1799 PTO ALEX  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 475 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; US-08-587-389-1

Query Match 92.1%; Score 2686; DB 1; Length 475;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-182;  
 Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLGVLVLGALAGLGPAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRHLH 60

Db	1	MLGVVLGALALAGLGFPAPEPQPGSGOCVEHDCFALYPGPATFLNASQICDGLRGHLM	60
Qy	61	TVRSSVAADVISLLNGDGVGRRRLWIGLQPPGCGDKRLGRLRGFOWVTGDNNTSYS	120
Db	61	TVRSSVAADVISLLNGDGVGRRRLWIGLQPPGCGDKRLGRLRGFOWVTGDNNTSYS	120
Qy	121	RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEPHFPATCRPLAV	180
Db	121	RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEPHFPATCRPLAV	180
Qy	181	EPGAAAAAVSITGTTPFAARGADFQALPYGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP	240
Db	181	EPGAAAAAVSITGTTPFAARGADFQALPYGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP	240
Qy	241	GAWDCSVENGGBHACNAIPGAPRCOCPAGALQADGRSCTASATQSCNDLCEHFCVNP	300
Db	241	GAWDCSVENGGBHACNAIPGAPRCOCPAGALQADGRSCTASATQSCNDLCEHFCVNP	300
Qy	301	DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYPNYDLVDG	360
Db	301	DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYPNYDLVDG	360
Qy	361	ECVEPVDPCFRANCEYQCQPLNQTSLVCVCAEGFADIPHEPHRCQMFNCQTACPADCDPN	420
Db	361	ECVEPVDPCFRANCEYQCQPLNQTSLVCVCAEGFADIPHEPHRCQMFNCQTACPADCDPN	420
Qy	421	TOASCECEGYIILDDGFICTDIDECENGFGCSGVCHNLPGTFCICGPDALARH	475
Db	421	TOASCECEGYIILDDGFICTDIDECENGFGCSGVCHNLPGTFCICGPDALVRH	475

Search completed: December 16, 2002, 17:27:18  
Job time : 15.5 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:26:23 ; Search time 8.5 Seconds  
(without alignments)  
1011.502 Million cell updates/sec

Title: US-09-509-994-2  
Perfect score: 2916  
Sequence: 1 MLGVLVLGALALAGLGFAP.....PSPTPGSTLTTPPAVLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105981 seqs, 1662342 residues

Total number of hits satisfying chosen parameters: 105981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2912	99.9	575	10	US-09-938-405-2
2	585	20.1	652	10	US-09-789-919-96
3	547	18.8	644	10	US-09-789-919-62
4	547	18.8	644	10	US-09-789-919-73
5	353	12.1	776	9	US-10-000-512-8
6	353	12.1	959	9	US-10-000-512-10
7	352.5	12.1	877	10	US-09-764-898-200
8	348.5	12.0	915	9	US-09-905-291A-34
9	348.5	12.0	915	10	US-09-909-320-34
10	348.5	12.0	915	10	US-09-909-088B-34
11	332	11.4	999	10	US-09-747-371-2
12	327.5	11.2	974	10	US-09-898-570-14
13	327.5	11.2	1009	10	US-09-898-570-16
14	325	11.1	800	10	US-09-764-853-800
15	325	11.1	800	10	US-09-764-898-272
16	323	11.1	534	10	US-09-804-156-14
17	323	11.1	534	10	US-09-946-633-6
18	314	10.8	997	10	US-09-747-371-3
19	313.5	10.8	845	10	US-09-898-570-12

20	311	10.7	433	10	US-09-925-302-691	Sequence 691, App
21	308	10.6	448	9	US-10-041-016-2	Sequence 2, Appli
22	308	10.6	448	10	US-09-083-002-2	Sequence 2, Appli
23	308	10.6	448	10	US-09-275-805-2	Sequence 2, Appli
24	308	10.6	448	10	US-09-836-561-1	Sequence 1, Appli
25	302	10.4	448	9	US-10-056-500-15	Sequence 15, Appl
26	302	10.4	448	12	US-10-052-586-408	Sequence 408, App
27	289.5	9.9	493	9	US-09-974-298-52	Sequence 52, Appl
28	284	9.7	735	10	US-09-898-570-10	Sequence 10, Appl
29	279.5	9.6	387	10	US-09-836-561-5	Sequence 5, Appli
30	278	9.5	368	10	US-09-275-805-9	Sequence 9, Appli
31	272	9.3	2444	10	US-09-944-849-2	Sequence 2, Appli
32	269	9.2	1055	10	US-09-855-722-2	Sequence 2, Appli
33	268	9.2	1212	10	US-09-855-722-3	Sequence 3, Appli
34	268	9.2	1238	10	US-09-855-722-5	Sequence 5, Appli
35	268	9.2	1238	10	US-09-944-849-4	Sequence 4, Appli
36	266.5	9.1	509	9	US-09-905-291A-315	Sequence 315, App
37	266.5	9.1	509	10	US-09-909-320-315	Sequence 315, App
38	266.5	9.1	509	10	US-09-909-088B-315	Sequence 315, App
39	266.5	9.1	509	12	US-10-052-586-52	Sequence 52, Appl
40	251.5	8.6	816	10	US-09-976-165-37	Sequence 37, Appl
41	240	8.2	886	10	US-09-992-647-11	Sequence 11, Appl
42	237.5	8.1	362	10	US-09-764-898-226	Sequence 226, App
43	234.5	8.0	1404	10	US-09-944-849-8	Sequence 8, Appli
44	234	8.0	722	10	US-09-908-322-12	Sequence 12, Appl
45	230.5	7.9	1246	10	US-09-919-497-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1  
US-09-938-405-2  
; Sequence 2, Application US/09938405  
; Patent No. US20020111296A1  
; GENERAL INFORMATION:  
; APPLICANT: Festoff, Barry W.  
; APPLICANT: Morser, Michael J.  
; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Inju  
; FILE REFERENCE: 51960AUSMI  
; CURRENT APPLICATION NUMBER: US/09/938,405  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: 60/229,714  
; PRIOR FILING DATE: 2000-08-31  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 2  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-938-405-2

Query Match	99.9%	Score	2912;	DB	10;	Length	575;
Best Local Similarity	99.8%	Pred. No.	1.3e-174;				
Matches	515;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps	0;						
Qy	1	MLGVLVLGALALAGLGFAPAP	PQPGSGCQVHDCFCALYPGPATFLNASQICDGLRGLHM	60			
Db	1	MLGVLVLGALALAGLGFAPAP	PQPGSGCQVHDCFCALYPGPATFLNASQICDGLRGLHM	60			
Qy	61	TYRSSVAADVISLLNGDGGVRRRLWIGLQPPGGCDPKRLGLRGFQWVTGDNNTSYS	120				
Db	61	TYRSSVAADVISLLNGDGGVRRRLWIGLQPPGGCDPKRLGLRGFQWVTGDNNTSYS	120				
Qy	121	RWARDLNCAPLCGPLICVAVSAEATVPSEPIWEEQOCVKADGFCFHFATCRPLAV	180				
Db	121	RWARDLNCAPLCGPLICVAVSAEATVPSEPIWEEQOCVKADGFCFHFATCRPLAV	180				
Qy	181	EPGAAAAAIVSYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP	240				
Db	181	EPGAAAAAIVSYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP	240				
Qy	241	GAWDCSVENGCGEHACNAIPGAPRCQCQAGAAALQADGRSCTASATQSCNDLCEHFCVPNP	300				

Db 241 GAWDCSVENGCEHACNAIPCAPRCQCPAGAAQADGSCSTASATQSCNDLCEHFCVNP 300  
QY 301 DQPSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
Db 301 DQPSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
QY 361 ECVBPVPCFRANCEYOCPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCPN 420  
Db 361 ECVBPVPCFRANCEYOCPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCPN 420  
QY 421 QAACECEGYILDDGICTDIDCEGNGFCGVCCHNLPGTFECICGPDSSALARIHGTDC 480  
Db 421 QAACECEGYILDDGICTDIDCEGNGFCGVCCHNLPGTFECICGPDSSALARIHGTDC 480  
QY 481 DSGKVDGDSGSGRPSPPTPGSTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGRPSPPTPGSTLTTPPAVGLVHSG 516  
RESULT 2  
US-09-789-919-96  
; Sequence 96, Application US/09789919  
; Patent No. US20020064855A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor  
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPIETIC BLOOD FORMING STEM  
; FILE REFERENCE: 2275-1-005  
; CURRENT APPLICATION NUMBER: US/09/789,919  
; CURRENT FILING DATE: 2001-02-21  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 96  
; LENGTH: 652  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-789-919-96

Query Match 20.1%; Score 585; DB 10; Length 652;  
Best Local Similarity 31.8%; Pred. No. 1.3e-29;  
Matches 181; Conservative 56; Mismatches 209; Indels 124; Gaps 32;  
QY 2 LGVLVLCALALA--GLGEPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLH 59  
Db 5 MGLLLLLLLLLLTPQAGTGADTE----AVVCGTACTYAHSGKLSAAEAQNHNCNGNGL 60  
QY 60 MTRVSSVAAD----VISLLINGDGGVGR--RLWIGLQLPPG--CGDPKRLGLRGLFQWVT 112  
Db 61 ATVKSEEAHQVORVLAQLLRREAALTARMSKFWIGLQREKKGKCLDPSL--PLKGFSSWG 118  
QY 113 GONNTYSRWARLDLNGAPLCPGLCVA--VSAAEATVPSE--PIWEEQOCEV-----EVRKAD 163  
Db 119 GGEDTYSNWHKELRNSC--ISKRCVSLDLDSQLPRLPNLPAWSGPGCGSPGCSGNTIE 176  
QY 164 GFLCEHFPATCRPLAY--EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQL 222  
Db 177 GFVKFSFSGMCRALGFGP-----QVTTYTTPFOTSSLSLEAVPFASAAVNA----- 224  
QY 223 MC-TAPPGAVQGHW---AREAPGMD-----CSVENGCCCHAC-NAIPGAP 263  
Db 225 -CGEGDKDTSQHYFLCKEKAADPDVWGSSGGLCVSPKYGCNFGNGGCHQDCFEFGDGSF 283  
QY 264 RCQCPAGAAQADGSCSTASATQSCNDLCE--HFCVNPNDQPCSYSCMCTGYRLAADQH 321  
Db 284 LCGRPGRELLDDLYTC-ASRNPSSSPCRGGATCVLGP-HGKNTCRCPQGYQLDSSQL 341  
QY 322 RCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGECVEPVDPCFRANCEYOCQPL 381  
Db 342 DCDVYDEC--QDSPCAQECVNTPGFRCEWVG-----EPGGP-----GEGACQDV 386  
QY 382 NCTSYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCPNQTAS--CECEGYIL--DDGF 437

Db 387 DE-----CALGRSP-----CAQGC--TNTDGSFHCSEEGVYLAGEDGT 423  
QY 438 ICTDIDEC--ENGFCGVCCHNLPGTFECICGPDSSALARIHGTDCDSKV-----D 486  
Db 424 QCQDVDECVGPGGLDCLCFNTQGSFHCGLPGLVWLAPN-GVSCMTGPFVSLGPPSGPDD 482  
QY 487 GGDGSGSG-----PPPSPTPGSTLTTPPA 509  
Db 483 EEDKGEREGSVTPRAATASPTGPEGTPKA 512  
RESULT 3  
US-09-789-919-62  
; Sequence 62, Application US/09789919  
; Patent No. US20020064855A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor  
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPIETIC BLOOD FORMING STEM  
; FILE REFERENCE: 2275-1-005  
; CURRENT APPLICATION NUMBER: US/09/789,919  
; CURRENT FILING DATE: 2001-02-21  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 62  
; LENGTH: 644  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-789-919-62

Query Match 18.8%; Score 547; DB 10; Length 644;  
Best Local Similarity 30.2%; Pred. No. 2.9e-27;  
Matches 168; Conservative 55; Mismatches 210; Indels 124; Gaps 27;  
QY 5 LVLGALALAGLGPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRS 64  
Db 9 LLIGLGLQGPWAG--AAADSO--AVVCEGTACTYAHWGKLSAAEAQHRCEGNGNLATVKS 64  
QY 65 SVAA-----DVISLLINGDGGVGR--RLWIGLQLPPGCPKRLGLRGLFQWVTGDNNTS 118  
Db 65 EEARHVQQAQALTLKTKAPLEAKMGKFWIGLQREKNCNTYHDL-PMRGFSWVGGEEDTA 123  
QY 119 YSRWARLDLNGAPLCPGLCVAVSAAEATVPSE--PIWEEQOCEV-----KADGFLCEHF 171  
Db 124 YSNWYKASKSSCIFKRCVSLILDLTLTPHPSHLPKWHESPCGTPEAPGNSIEGFLCKFNF 183  
QY 172 PATCRPLAY--EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQ-----LM 223  
Db 184 KGMCRPLALGGPG-----RVTTYTTPFQATSSLEAVPFASVANVA--CGDEAKSETHYFL 236  
QY 224 CT-APGAVQGHWAEP---CAWDCSVENGCCCHAC-NAIPGAPRCQCPAGAAQADG 277  
Db 237 CNEKTEGIF--HMGSSGGLCVSPKFGCNGNGGQDCFEFGDGSFRCGCRPFLRDL 294  
QY 278 RSCATASQSCNDLCEHFCVNPNDQPG-----SYSCMCTGYRLAADQHRCEVD 327  
Db 295 VTCAS-----RNPSSNPTCGGMCHSVPLSENITCRCPSGYQLDSSQVHCVDID 344  
QY 328 DCILEPSPQRCVNTQGGFECHCYPNYDLVDG---ECVEPVDPCFRANCEYOCQPLNQT 384  
Db 345 EC--QDSPCAQECVNTLGSFHCCEWVGQ-PSGPKREACEDVDECAANSP----- 392  
QY 385 SYLCVCAEGFAPIPHEPHRCQMFNCNTACPADCPNQTAS--CECEGYIL--DDGFICT 440  
Db 393 -----CAQGC-----NTDGSFYCSCKEGYIVSGEDSTQCE 423  
QY 441 DIDECEEN--GGFCGVCCHNLPGTFECICGPDSSALARIHGTDCDSG-----KVD 486  
Db 424 DIDECSARGNPCDLSLFTNTDGSFRCGCPPGWELAPN-GVFCSRGTVFSELPARPQKED 482  
QY 487 GGDGSGSGEPPSPPTGSG 503









APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
ACIDS  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: 10466-14  
CURRENT FILING DATE: 1999-07-07  
CURRENT FILING DATE: 2001-07-18  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 34  
LENGTH: 915  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-909-088B-34

Query Match 12.0%; Score 348.5; DB 10; Length 915;  
Best Local Similarity 34.3%; Pred. No. 9.5e-15;  
Matches 85; Conservative 36; Mismatches 108; Indels 19; Gaps 14;

QY 245 CSVNGGCEACNAIPGAPRCQCPAGAAQADGRSCTA-SATQSCNDLCEHFVCPVNPDPQ 303  
DB 283 CAMEDHNCQELCVNVPQSFVQCYSGYALAEADCKRCVAVDYCASENHGCHECV-NAD-- 339

QY 304 GSYSCMETGYRLAADHRCEDVDCCILEPSPQRCVNTQGGFECHCYPNYDL-VDGEC 362  
DB 340 GSYLCQCEHGFALNDEKTRINICALNKPGEHCVNMEESYICRHRGYTLDPNGKT 399

QY 363 VEPVDCFRAN--CEYQCPN-OTSILVCVCAEGFADIPHEPHRCOM--FC--NOTACPA 415  
DB 400 CSRDVHCAQDQHGCEJLC--LNTEDSFVQCSEGLF-INEDLTKTSRVDYCLLSDHGCEY 456

QY 416 DC-DPNTQASCECEGYIL-DDGFICTDIDECENGGF-CSGVCHNLPPTGTECICGPDPSAL 472  
DB 457 SCVNDRSFACQCEGHVLRSDGKTCALDSCALGDHGCHEHSCVSSSDSFVQCQCF-EGYI 515

QY 473 ARHGTDC 480  
DB 516 LREDGKTC 523

RESULT 11  
US-09-747-371-2  
Sequence 2, Application US/09747371  
Patent No. US20020006616A1  
GENERAL INFORMATION:  
APPLICANT: Gish, Kurt  
APPLICANT: Mack, David  
TITLE OF INVENTION: No. US20020006616A1 Methods of Diagnosing Breast Cancer, Com  
BIOLOGICAL  
TITLE OF INVENTION: Screening for Breast Cancer Modulators  
FILE REFERENCE: A-69028/DJB/JJD  
CURRENT APPLICATION NUMBER: US/09747371  
CURRENT FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: PCT/US/00/06952  
PRIOR FILING DATE: 2000-03-15  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 999  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-747-371-2

Query Match 11.4%; Score 332; DB 10; Length 999;  
Best Local Similarity 26.8%; Pred. No. 1.1e-13;  
Matches 101; Conservative 31; Mismatches 117; Indels 128; Gaps 18;

QY 182 PGAAAAVSITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWARE--- 238  
DB 9 PGAAWAVLLLLLLLP-----PL-LLLAGAVPPGGRGAAGQEDVD 47

QY 239 --APGAWDCSVENGCGEHA---CNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDL-- 291  
DB 48 ECAQGLDDC-----HADALCQNTPTSYKCSCKPG--YQEGEQC-EDIDECNGLNG 96

QY 292 -CEHFCVPNPDPGYSVSCMETGYRLAADHRCEDVDCCILEPSPQRCVNTQGGFECH 350  
DB 97 GCVHDC---NIPGNVRCCTGDFGFMHLADHNCGLDDECLNNGGCGQTCVNVMSYBCC 153

QY 351 CYPNYDLVDGE--CPEVD---PCFRAN--CEYQCPN-OTSILVCVCAEGF----- 394  
DB 154 CKEGFFLSDNQHTCIHRSEGLSCMKNKHGCHSHICKAPRGVACVACRPGFELAKNQDC 213

QY 395 -APIPHEPHRCOMFCNOTA----- 412  
DB 214 ILTCNHGNGCGQHSCDDTDADGPECSCHPQYKMHDTGRCCLEREDVLEVTESNTTSVVDG 273

QY 413 -----CPADC-DPNTQASCECEGYILD-DGFICTDIDEC--NGG 449  
DB 274 DKRVKRRLLMETCAVNNGGCDRTCKDTSTGVHSCSPVGFLLQDGTCKDIDECOTRNGG 333

QY 450 FCSGVCHNLPPTGTECIC 466  
DB 334 -CDHFCNKIVGSDGCG 349

RESULT 12  
US-09-898-570-14  
Sequence 14, Application US/09898570  
Patent No. US20020123612A1  
GENERAL INFORMATION:  
APPLICANT: GERLACH, VALERIE L.  
APPLICANT: ELLERMAN, KAREN  
APPLICANT: MACDOUGALL, JOHN R.  
APPLICANT: SMITHSON, GLENDA  
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND  
METHODS OF USING THE SAME  
FILE REFERENCE: 15966-776CIP  
CURRENT APPLICATION NUMBER: US/09/898,570  
CURRENT FILING DATE: 2001-07-03  
PRIOR APPLICATION NUMBER: 60/198,293  
PRIOR FILING DATE: 2000-04-19  
PRIOR APPLICATION NUMBER: 60/198,645

RESULT 14  
US-09-764-853-800  
; Sequence 800, Application US/09764853  
; Patent No. US20020090672A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```

RESULT 13
US-09-898-570-16
; Sequence 16, Application us/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776CIP

```

